THE AMERICAN
JOURNAL OF PSYCHOLOGY

Founded in 1887 by G. STANLEY HALL

Vol. LII JULY, 1939 No. 3

PAIN: HISTORY AND PRESENT STATUS

By KARL M. DALLENBACH, Cornell University

The doctrine of the five senses, which is common sense today, is ascribed to Aristotle, although it is probable that it was also common sense in his day and generation. However that may be, the doctrine descends to us from antiquity, and its does not include, nor does it make any provision for, a sense of pain. Though Aristotle states that touch, the fifth of the senses enumerated by him, "discriminates several sense-qualities,"[1] pain is explicitly not one of them; for he, like Plato before him, places pain with pleasure among the passions of the soul.

It is not surprising that the Ancients should have set pain apart from the senses, or that they should have considered it as the antithesis of pleasure. Unlike sight, hearing, smell, taste, and touch, pain does not possess an obvious sense-organ; nor is it restricted to any part of or locus in the body, but rather pervades the whole. Moreover, it is not a quality of external objects as the qualities of the five senses were supposed to be; and it does seem to be inherently unpleasant, as most pains (probably more than 99% of them—indeed, a recent writer, Fröbes,[2] says *all*) may thus be characterized. Once pain, identified with unpleasantness, is set in opposition to pleasure, it transcends the limitations of the sensory field and extends into the other fields of mental life.

These considerations, together with the dominant influence of Aristotle's pronouncements upon subsequent philosophical and scientific thought, long delayed the recognition of pain as a sensory quality. The affective classification of pain and the doctrine of the five senses prevailed until

* Presidential address of the Eastern Psychological Association, given at Bryn Mawr College, March 31, 1939.

[1] Aristotle, *Treatise on the Principles of Life*, Eng. trans., W. A. Hammond, 1902, bk. ii, chap. 6; bk. iii, chap. 1.

[2] Joseph Fröbes, *Lehrbuch der experimentellen Psychologie*, 1, 1923, 149.

BOSTON UNIVERSITY
COLLEGE OF LIBERAL ARTS
LIBRARY

comparatively recent times. All through the intervening years, however, some intrepid souls have considered and variously answered the question tacitly raised by Aristotle regarding touch: "Is it a bundle or group of many special senses?" Since pain was turned into a special sense as a result of the attempts to analyze touch, it may be profitable as well as interesting to review the answers that have been given to that question.

Aristotle's commentators were divided. Themistius [317-387] adopted the view that touch was a plurality of senses.[3] He held, though he did not specify, that there were as many separate senses in touch as there were different tactile qualities. Alexander, in the 2nd century, Simplicius, in the 6th, and Philoponus, in the 7th, were, however, of the contrary opinion,[4] and, since the 'master' had not definitely spoken, but had merely mooted the point, the contrary opinion that touch was a single sense prevailed for their day and generation.

The question was raised again in the 11th Century by the Schoolmen. Themistius' doctrine, that touch was a plurality of senses,[5] was adopted by Avicenna [980-1038] and Averroes [1126-1198] among the Arabian Schoolmen, and by Appollinaris [12th century], Albertus Magnus [1193-1280], Aegidius [1247-1316], Jandunus [d. 1328], and Marcellus [14th century] among the Latin Schoolmen. They were not, however, in agreement regarding the kind or the number of senses into which it was to be divided. Avicenna, like Themistius, divided it into as many senses as there were tactile qualities. He did not, however, specify their number, though he did distinguish two from touch proper; namely, the sense of pain from a wound and the sense of titillation.[6] Averroes, his compatriot, accepted titillation; but he disagreed regarding pain, and added hunger and thirst.[7] Aegidius increased the number of senses by two: one for temperature (hot and cold) and the other for humidity (dry and moist).[8]

Among the more modern philosophers, Cardano [1501-1576] divided touch into four senses: one for the tactile qualities of hot, cold, wet, and dry; a second for light and heavy; a third for pleasure and pain; and a fourth for titillation.[9] These were denied by Scaliger [1484-1558] who himself distinguished sexual appetite as a sixth sense. Francis Bacon[10] [1561-1626] distinguished sex and so also did Voltaire[11] [1694-1778] and Buffon[12] [1707-1788]. Locke [1632-1704] added hunger and thirst;[13] and Kant, [1724-1804] a vital sense (*sensus vagus*) into which he placed heat, cold, shudder, quake, thrill, hunger and thirst.[14]

Beginning in the 2nd century A.D., the physicians and scientists also tried their hand at analyzing touch. Galen [131-200] thought that in addition to touch proper there was a sense of heat and cold.[15] Pechlini, in 1691, divided the fifth sense into touch, cold, and heat.[16] In the 18th century, Erasmus Darwin, [1731-1802]—who

---

[3] Cited by William Hamilton, *Lectures on Metaphysics and Logic,* edited by H. L. Mansel and John Veitch, 2, 1870, 155.

[4] *Idem.* [5] *Idem.* [6] *Idem.*

[7] *Ibid.*, 156. [8] *Ibid.*, 155 f. [9] *Ibid.*, 156.

[10] *Idem.* [11] *Idem.* [12] *Idem.*

[13] *Idem.* [14] *Idem.* [15] *Idem.*

[16] J. N. Pechlini, *Observationum physico-medicarum libri tres,* 3, 1691, 410; quoted from Torsten Thunberg, Physiologie der Druck-, Temperatur- und Schmerzempfindungen, in W. Nagel's *Handbuch der Physiologie des Menschen,* 3, 1904, 648.

had the misfortune to be the grandfather of Charles Darwin, since otherwise he would have been great in his own right—added seven senses to the traditional five: the sense of heat, of muscular extension, of hunger, of thirst, of suffocation, of sexual appetite, and of lactiferousness.[17]

In support of his statement that "nature has provided us with a set of nerves for the perception of heat,"[18] Erasmus Darwin cites an interesting observation made on a patient by one of his relatives, Dr. R. W. Darwin. This patient, as a result of fever and violent cramps in his legs, had lost sensitivity to touch in one of his feet. "Though pricked with a pin in five or six places," Darwin reports, "the patient declared he did not feel it in the least, nor was he sensible of a very smart pinch. I then held a red hot poker at some distance, and brought it gradually nearer, till it came within three inches, when he asserted that he felt it quite distinctly. I suppose some violent irritation of the nerves of touch had caused the cramps, and had left them paralytic; while the nerves of heat, having suffered no increase of stimulus, retained their irritability."[19]

Some of Darwin's observations regarding the other senses that he proposed are also worthy of quotation. In reference to the sense of extension, he writes: "the whole set of muscles, whether they are hollow ones, as the heart arteries and intestines, or longitudinal ones attached to bones, . . . may be considered as one organ of sense, and the various attitudes of the body, as ideas belonging to this organ."[20] The sense of hunger, he writes, "is most probably perceived by those numerous ramifications of nerves that are seen about the upper opening of the stomach;" and thirst "by the nerves about the fauces and top of the gula."[21]

At the end of the section on the organs of sense, Darwin adds that "there are many more senses than have been here mentioned . . . which have not acquired the name of senses."[22] Pain, however, was not one of these, for he regarded pain and pleasure as central effects of sensorial motions. Pain results, he says, "whenever the sensorial motions are stronger than usual. . . . A great excess of light . . . of pressure or distension . . . of heat . . . of cold produces pain."[23] He thus anticipates the intensive theory of pain that was introduced several generations later.

About a decade after Darwin, Bell [1774-1842] posited,[24] without reference to earlier writers, a sixth sense—a sense of movement—in addition to the traditional five. This new sense, whose organs are the muscles of the body, was, he believed, "the most important of all . . . for it is by a sense of motion that we know many of the qualities of outward things, as their distance, shape, resistance, and weight[25] . . . hardness, softness, figure, extension, and motion."[26]

Bell also writes of hunger and thirst as sensations. "We are," he states, "solicited to take food by the uneasy sensation of hunger," which is "the effect of the attrition of the sensible coats of the stomach upon each other by the peristaltic motion of the stomach and compression of the viscera."[27] "Thirst is a sensation seated in the

[17] Erasmus Darwin, *Zoonomia*, 1, 1794, 76-90, sec. xiv.
[18] *Ibid.*, 86. [19] *Ibid.*, 87. [20] *Ibid.*, 87.
[21] *Ibid.*, 89. [22] *Ibid.*, 90. [23] *Ibid.*, 90.
[24] Charles Bell, *Anatomy of the Human Body*, 3, 1803, 224-372; John and Charles Bell, *The Anatomy and Physiology of the Human Body*, 5th Amer. ed., revised by Charles Bell, 2, 1827, 154-290, esp. 158, 290.
[25] John and Charles Bell, *op. cit.*, 158.
[26] *Ibid.*, 290.
[27] *Ibid.*, 326.

tongue, fauces, esophagus, and stomach. It depends on the state of the secretions which bedew these parts and arises either from a deficiency of secretion or from an unusual acrid state of it."[28] He took no account of either of them, however, in his classification of the senses; nor did he specifically mention pain.

Approximately 100 years ago, in 1840, Johannes Müller [1801-1858] presented his theory of the specific energies of nerves. Though this theory was anticipated by Bell, as has frequently been pointd out, it was independently formulated by Müller and its scientific significance was not recognized until Müller had published it. In accordance with the traditional doctrine of the senses, Müller posited five kinds of sensory nerves.[29] The fifth class, the nerves of feeling (*Gefühlsnerven*), yielded, he believed, a number of different sensations: tickle, itch, shudder, pleasure, pain, fatigue, suffocation, warmth, cold, touch, and movement. He did not, however, posit different nerve-specificities for those qualities but attempted to account for them by differences in their mode of arousal and in the state of the organism at the time of their arousal. There is, therefore, as far as the analysis and classification of the senses are concerned, little in Müller beyond Aristotle. Müller is, nevertheless, important to us for two reasons: first, because his doctrine of nerve-specificity effectively destroyed, in the realm of scientific thought at least, the point of view, against which the philosophers have inveighed since Descartes and Locke, that sensory nerves conduct to the brain properties or incorporeal copies of the objects perceived. He showed definitely that that was not the case, as different objects produce the same effect with the same nerves, and the same object produces very different effects with different nerves. He is important to us, secondly, because his doctrine provided a sanction for analysis and a systematic basis for classification. If a given nerve, however affected, gives rise to a sense quality that depends upon the specific character of that nerve, may it not be then that the various nerves giving rise to different qualities of touch or feeling have different specificities? The question raised by Aristotle has now obtained systematic significance.

Before Müller, there was no unifying doctrine to shape analysis nor any accepted principle of classification. The efforts of the various inquirers who have been mentioned were without lasting significance. Their individual insights were merely eddies in the current of thought. The gains made by them were not consolidated. Every author wrote, against the background of the traditional doctrine of the five senses, as if he were pro-

[28] *Ibid.*, 327.
[29] Johannes Müller, *Handbuch der Physiologie des Menschen*, 2, 1840, 249-502.

posing something new. Lapses of centuries and differences in language may in part account for the lack of continuity, but even Bell, Darwin's contemporary in England, wrote of the 'muscle sense' as if he were discussing something new. He gave no indication of knowing Darwin's work.

The testimony of the ages proved ineffectual in refuting the Aristotelian doctrine of the senses. Negative evidence usually is ineffectual. Doctrines pass from the scientific stage, not because they have been discredited but because they have been superseded—pushed off the stage and replaced by others. What had been lacking all along to consolidate the gains as they were being made was a positive principle; and Müller, though he himself, as we have seen, accepted the traditional doctrine of the five senses, gave us such a principle. His work may consequently be reckoned as the dividing line between the ancient and the modern history of the psychology of the senses.

Under the impetus of the Müllerian doctrine, scientific thought was tremendously quickened. The principle of the dependence of qualitative difference upon specific nerves, though denied by some,[30] was immediately extended by others to the different sense departments, and Müller's "sense of feeling," which is of particular interest to us, was divided into several senses.

For example, in 1844, Natanson divided the sense of feeling into touch, temperature, and resistance.[31] Weber, in 1846, divided it into touch and common sensibility (*Gemeingefühl*).[32] Touch was restricted by Weber to the skin; and common sensibility, which stems from Kant's *"sensus vagus,"* was defined to include all the qualities common to the skin and the other organs and tissues of the body. These two classes were then further divided. "Touch," he wrote, "provides us with two kinds of sensations which are peculiar to it, pressure sensations (*Druckempfindungen*) and temperature sensations (*Temperaturempfindungen*); at the same time the organ of touch and its nerves are so constituted that the same sensations may be differentiated from one another when they are aroused at two different places on the skin. We may, therefore, distinguish the sense of locality, the sense of pressure, and the sense of temperature."[33] These senses were restricted to the skin because he had failed to elicit them from the internal organs and deep tissues exposed in operations and wounds.

The sensations of common sensibility (*Gemeingefühlsempfindungen*) were even

[30] Cf. R. H. Lotze, *Allgemeine Pathologie und Therapie,* 1842, 164; also A. W. Volkmann, Von der specifischen Reizbarkeit der Nerven, in Rudolph Wagner's *Handwörterbuch der Physiologie,* 2, 1844, 521-526.

[31] — Natanson, Analyse der Funktionen des Nervensystems, *Arch. f. physiol. Heilkunde,* 3, 1844, 515-535, esp. 517 f.

[32] E. H. Weber, Der Tastsinn und das Gemeingefühl, in Rudolph Wagner's *Handwörterbuch der Physiologie,* 3, 1846, 481-588. Reprinted separately as *Die Lehre von Tastsinne und Gemeingefühl,* 1851.

[33] *Op. cit.,* 511.

more varied in kind. They derive, as he thought, from all parts of the body possessing nerves (the bones, cartilage, tendons, enamel of the teeth, nails, hair, and epidermis being the exceptions); and are characterized by the fact that "we cannot relate them to outer objects."[84] From internal stimuli (such as chemical and nutritive changes in muscles, organs, and blood) derive fatigue, hunger, thirst, nausea, feelings of well-being and of sickness, etc. From external stimuli and the movement of muscles derive pain, sensations of movement, shudder, tickle, itch, sting, etc. The category of *Gemeingefühl* is Weber's scrap basket into which he cast a miscellany of sensory remnants. It is, as Henle [1809-1885] wrote of it, "the sum, the unsorted chaos of sensations from all the sensitive parts of the body, which leads to the consciousness of self."[85]

Common sensibility, as Weber believed, was finest in the muscles and skin, and pain was its most marked characteristic. Like Darwin, Romberg, Henle, and Volkmann before him,[86] he anticipated the intensive theory of pain, as he wrote that pain "is aroused in all organs of sense by external stimuli which are so strong that they affect not only the ends of the nerves, but also their trunks."[87]

From Weber on down to the present, numerous analyses of the sense of touch have been made and numerous classifications proposed. In all of them, *Gemeingefühl*—or *coenaesthesis*, as Weber also called it—has played an important though varying rôle, serving ever as the classificatory term for the residue of unsorted sensations and feelings.

The divers classifications are, however, no longer of concern to us. The principle of analysis having been accepted, we may now turn to our special interest, pain, and follow its rapid unraveling—and snarling—from the skein of theory.

Several investigators, as we have seen, anticpated the intensive theory of pain. It was first explicitly formulated, however, by Erb, in 1874,[88] who held that pain was not a sensation specific in kind, but one of high degree. Every sensory stimulus, he maintained, was capable of producing pain if it reached sufficient intensity. This theory, as we shall see, was accepted and supported by many writers.

The theory that pain was a separate and distinct sense, first suggested by Avicenna in the 11th century, was intimated for a second time by Lotze in 1852.[89] It was first definitely formulated, however, by Schiff in 1858, following his analgesic experiments on animals.[90] Noting the

---

[84] *Ibid.*, 495.
[85] Henle, *Allgemeine Anatomie*, 1841, 728; quoted from Weber, *op. cit.*, 563.
[86] Volkmann, *op. cit.*, 520.
[87] Weber, *op. cit.*, 495.
[88] Erb, *Krankheiten der peripherischen cerebrospinalen Nerven*, 1874. Quoted from G. W. A. Luckey, Some recent studies of pain, this JOURNAL, 7, 1895, 109.
[89] R. H. Lotze, *Medicinische Psychologie*, 1872, 247.
[90] Schiff, *Lehrbuch der Physiologie*, 1, 1858, 228. (Cf. O. Funke, Der Tastsinn und die Gemeingefühle, in L. Hermann's *Handbuch der Physiologie der Sinnesorgane*, 3, (2), 1879, 297.

effects of various incisions in the spinal cord, he found that pain and touch were independent. When he cut through the gray matter, pain could not be aroused below the level of the incision while touch was left intact; when he cut through the white matter, but not the gray, touch was lost but pain remained. The results of these vivisections were immediately corroborated by clinical evidence, as physicians reported numerous pathological cases of diseased or injured spinal cords with similar sensory defects.[41] Cases of tabes dorsalis with a marked differentiation in the time of arousal of touch and pain—the pressure from the jab of a needle preceding the pain by several seconds—were also placed in evidence.

This theory of pain, which may be designated as the sensory theory, was reaffirmed by Funke in 1879 after a thorough examination of the evidence pro and con.[42] It did not, however, come to full bloom until after the classical experiments of Blix and Goldscheider in the early '80s. Blix,[43] in 1882, and Goldscheider,[44] in 1884,—in experiments undertaken to extend the Müllerian doctrine of specific nerve-energies to touch —independently discovered in the skin separate spots for warmth, cold, pressure, and pain, which—no matter how stimulated—yielded, if stimulated, their own specific quality. These results were accepted by many as conclusive evidence for the theory that pain is a separate department of sense.

The intensive theory and the sensory theory stood in opposition to one another, and they both stood in opposition to the traditional pleasure-pain theory which represented pain as an affective *quale*. In the decade between 1886-1896, these different opinions clashed and a three-cornered controversy ensued, the like of which has never before, nor since, appeared in the scientific literature. The proponents of the contending views had to defend themselves from two directions and at the same time lead a divided attack. The controversy was intense; no quarter was asked and

---

[41] Cf. O. Funke, *op. cit.*, 299.

[42] Funke, *op. cit.*, 289-309.

[43] Magnus Blix, Experimentelle Beiträge zur Losung der Frage über die specifische Energie der Hautnerven, *Zsch. f. Biol.*, 20, 1884, 141 ff., 160 ff. This was a translation of a paper read in Sweden in November 1882 and first published in *Upsala läkareförens. förhandl.*, 18, 1883, 87 ff.

[44] Alfred Goldscheider, Die specifische Energie der Temperaturnerven, *Monatshft. f. prak. Derm.*, 3, 1884, 198-208, 225-241; Die specifische Energie der Gefühlsnerven der Haut, *ibid.*, 283; Die specifischen Functionen der Nerven der Haut, *Proc. 8th Internat. Cong. Kopenhagen*, 3, 1884, 25-27, also *Vtjsch. f. Derm., Wien*, 11, 1884, 313-316; Nachtrag zu den Mittheilungen über die specifischen Energieen der Hautnerven, *Monatshft. f. prac. Derm.*, 4, 1885, 5; Ueber Wärme-, Kälte- und Druckpunkte, *Verhandl. d. Physiol. Gesellsch., Berlin*, 1885, 000. All of these studies may be found in Goldscheider's *Gesammelte Abhandlungen*, 1, 1898, 53-107.

none given. Writers were dogmatic and caustic, offering the apology that this manner "made for brevity and explicitness of exposition."[45]

The traditional point of view, that pain was an affective *quale*, was supported by the philosophers and by psychologists of philosophical bent: notably by Hurwicz, Lehmann, and Wundt in Germany; by Bain, Bradley, Spencer, Sully, Stout, and Ward in England; and by Baldwin, Dewey, James, and Marshall in America. Of all of these, the American, Henry Rutgers Marshall, was the most active and articulate. He wrote one book and numerous articles in behalf of the theory[46] and though, as it turned out, he fought a losing battle, he fought determinedly and vigorously. The "weight of authority" was with him, as he himself pointed out; but that proved of slight advantage as the authorities were widely divided among themselves. There was not *one* pleasure-pain theory but many; indeed, there were almost as many different theories as there were writers on the subject; and, as a rule, the theories proposed were satisfactory only to the men proposing them. Concerted effort and the maintenance of a common front were, under those conditions, impossible. It was almost a case, among the pleasure-pain theorists, of every man for himself. Though Marshall did yeoman's service in behalf of the traditional point of view, the particular theory that he advocated was as unsatisfactory to most of the members of his own camp as it was to his opponents in the other two camps. The internecine quarrels among the pleasure-pain theorists probably did more to advance the theories of the opposition than all their criticisms did to check them.

The situation was very different in the opposing camps. The intensive and sensory theories were supported by physiologists and by younger psychologists with psychophysical interests. These men were untrammeled by tradition; they put their trust in experiment and observation, and followed where their results led, even though this meant, as it frequently did, change of allegiance. For example, when Blix discovered the pain spots in 1882, he was inclined to believe that pain was specific; but by 1885 later experiments moved him to conclude that it was not. "There

---

[45] Herbert Nichols, The origin of pleasure and pain, *Philos. Rev.*, 1, 1892, 404.

[46] H. R. Marshall, The classification of pleasure and pain, *Mind*, 14, 1889, 511-536; The physical basis of pleasure and pain, *ibid.*, 16, 1891, 327-354, 470-497; Pleasure-pain and sensation, *Philos. Rev.*, 1, 1892, 625-648; *Pain, Pleasure, and Aesthetics*, 1894, 1-364; Pleasure-pain, *Mind*, 19, 1894, 533-535; Are there special nerves for pain?, *J. Nerv. & Ment. Dis.*, 21, 1894, 71-84; Pleasure and pain, *Proc. Amer. Psychol. Ass.*, 1894, 24; Pleasure-pain and emotion, *Psychol. Rev.*, 2, 1895, 57-64; Emotion versus pleasure-pain, *ibid.*, 166-168; *Mind*, 4, 1895, 180-194; Physical pain, *Psychol. Rev.*, 329-347.

are," he writes in his second study, "three specific kinds of nerve apparatus in the skin, one for warmth, one for cold, and one for pressure. No specific organs for the sense of pain have been proved to be in the skin."[47] Goldscheider also shifted front. At first he did not believe that pain was specific; but by 1885, just when Blix was shifting in the other direction, he came to the conclusion that the evidence did speak in favor of specificity.[48] He held to that view until about 1891,[49] when he shifted once more—this time to the theory, which he fully developed in 1894,[50] that pain was mediated by the tactile nerves and resulted from an intensive summation of their excitations. He turned to that theory because of an experiment of Naunyn's,[51] in 1889, in which pain was aroused through the summation of weak stimuli rapidly applied. Naunyn found, in cases of tabes dorsalis, that a mechanical stimulus (a hair for example) that was below the threshold for touch or pain, when applied repeatedly from 60-600 times a second, yielded in a few seconds (from 6-20) a pain which soon became unbearable. Since he obtained similar results with other stimuli (electrical for example), he concluded that pain was the result of summation. This conclusion and Goldscheider's extension of it must be counted as variants of the intensive theory. According to that theory as originally formulated, pain arises from the intense stimulation of any sensory organ. Bright lights, loud sounds, strong smells and tastes, severe pressures, and extremes of temperature, all, it was thought, arouse pain if they are intense enough. The ultimate extension of this doctrine is exemplified in the writings of Külpe and Titchener. Külpe wrote in his *Grundriss* in 1893: "pain has been held to be a special quality of cutaneous sensation, and pain nerves put alongside of the nerves subserving heat, cold, and pressure. The hypothesis has found its chief support among physiologists. We cannot accept it, as it stands: for pain is produced in all cases where the stimulation of a sensory nerve passes a certain limit of intensity."[52] Titchener wrote in the first edition of his *Outline* in 1896:

[47] Blix, *op. cit.*, *Zsch. f. Biol.*, 21, 1885, 160 ff.

[48] Goldscheider, Neue Thatsachen über die Hautsinnesnerven, *Arch. f. Anat. u. Physiol.*, 1885, Physiol. Abth., Suppl. Bd., 1885, 1-110 (*Gesammelte Abhandlungen*, 1, 1898, 197 ff.).

[49] Goldscheider, Ueber die Summation von Hautreizen, *Arch. f. Physiol.*, 1891, 164-169 (*Gesammelte Abhandlungen*, 1, 1898, 397-432).

[50] Goldscheider, *Ueber den Schmerz in physiologischer und klinischer Hinsicht*, 1894, 1-66.

[51] B. Naunyn, Ueber die Auslösung von Schmerzempfindung durch Summation sich zeitlich folgender sensibler Erregungen, *Arch. f. exper. Pathol. u. Pharm.*, 25, 1889, 272-305.

[52] Oswald Külpe, *Outlines of Psychology*, 1893; Eng. trans., E. B. Titchener, 1905, 90.

"excessive stimulation of any sense-organ, or direct injury to any sensory nerve, occasions the common sensation of pain."[53]

The sensory theory, that pain is a sensation mediated by specific nerves, found its chief support, as Külpe remarked, among the physiologists. Blix and Goldscheider were among the pioneers holding this view. Though they later shifted position, as we have seen, their places were taken by others—notable among whom was Max von Frey. This theory, it should also be noted, received almost unanimously the support of the clinicians, the most prominent among them being S. Wier Mitchell,[54] in America, and Henry Head,[55] in England. Psychologists were long in joining this camp. The first and for a time the only psychologist to uphold this point of view was Herbert Nichols[56]—a former student of G. Stanley Hall and a man who had come to psychology with training in engineering and sixteen years of successful service with the Pennsylvania Railroad.[57] Unprejudiced and not indoctrinated with psychological lore, he was possessed of sufficient self-confidence to espouse and to take a leading rôle in a minority movement. He opined, that "the prejudice of science . . . ran in favor of specific end-organs,"[58] and, as it turned out, he was right; for, when Head published his clinical observations[59] and Von Frey his experimental conclusions[60] correlating warmth with Ruffini cylinders, cold with Krause end-bulbs, pressure with hair follicles and Meissner corpuscles, and pain with free nerve endings,[61] psychologists moved over to the sensory view almost in a body. Here was something definite to teach, and textbooks changed over night. For example: Titchener, who had written in 1896 from the point of view of the intensive theory, wrote in his *Primer* in 1898 "of the sensation of pain,"[62] and placed it alongside

[53] E. B. Titchener, *An Outline of Psychology*, 1896, 65.

[54] S. W. Mitchell, Precision in the treatment of chronic diseases, *Med. Rec.*, 42, 1892, 721-726; Post-hemiplegic pain, *Med. News*, 62, 1893, 421-423, Wrong reference of sensations of pain, *ibid.*, 66, 1895, 281 f.

[55] Henry Head, On disturbances of sensation with especial reference to the pain of visceral disease, *Brain*, 16, 1893, 1-133; 17, 1894, 339-480; 19, 1896, 153-276.

[56] Herbert Nichols, The origin of pleasure and pain, *Philos. Rev.*, 1, 1892, 402-432, 518-534; Pain nerves, *Psychol. Rev.*, 2, 1895, 487-490; The feelings, *Philos. Rev.*, 4, 1895, 506-530; Pain nerves, *Psychol. Rev.*, 3, 1896, 309-313.

[57] K. M. Dallenbach, Herbert Nichols: 1852-1936, this JOURNAL, 49, 1936, 320-321.

[58] Nichols, Pain nerves, *Psychol. Rev.* 3, 1896, 311.

[59] Head, *opp. citt.*

[60] Max von Frey, *Die Gefühle und ihr Verhältnis zu den Empfindungen*, 1894, 1-24; Beiträge zur Physiologie des Schmerzsinnes, *Ber. ü. d. Verhandl. d. k. sächs. Ges. d. Wiss. z. Leipzig*, math.-phys, Kl., 46, 1894, 185-196, 288-296; Beiträge zur Sinnesphysiologie der Haut, *ibid.*, 47, 1895, 166-184; 48, 1896. 462-468; Untersuchungen über die Sinnesfunctionen der menschlichen Haut: I. Druckempfindungen und Schmerz, *ibid.*, 49, 1897, 169-266.

[61] *Op. cit., ibid.*, 47, 1895, 181-184.

[62] Titchener, *A Primer of Psychology*, 1898, 45.

of the sensations of pressure, heat and cold; and in his *Instructor's Manual* he said: "investigation of the cutaneous sensations has moved so rapidly during the past five years that there is no adequate account of them to be found in the textbooks."[68]

Evidence for the sensory theory extant in 1896 was derived from many sources. *(1)* A large share was drawn from pathology, as clinical evidence indicated that the skin possessed four distinct forms of sensibility: warmth, cold, pressure, and pain; and that any one or any combination of them may be lost or impaired without involving the others.

(2) Experiments with intense stimuli also showed that pain was specific; for, contrary to the views of the intensive theorists, excessive stimulation of the special senses does not yield pain when the effect is limited to the organ concerned. For example: if the retina is stimulated, however intensely, it yields not pain but light. If proximal regions are stimulated simultaneously, pain results, but only because those extra-retinal regions are themselves capable of yielding pain. As with vision so also with audition, smell and taste. Excessive stimulation of any of the special senses does not yield pain unless the effects involve tissues that are themselves supplied with pain nerves. Even the warm and cold spots were found to be specific. No matter how excessive their stimulation—whether they are prodded with a probe, pricked with a needle, or burned with a red hot wire—they yield, when stimulation is restricted to the proper area, not pain, but their own specific quality.

(3) The long latent period, the comparatively slow rise of pain after stimulation, which is exaggerated in cases of tabes dorsales, was set in evidence of specificity.

(4) So also were the results obtained with anesthetics. Under light chloroform anesthesis, the patient is sensitive to the slightest touch, but to pain he is totally insensitive. Again, if an ice pack is applied to the elbow, pressure, warm, and cold, and finally pain lapse from the lower arm in that order. Similar results were obtained with cocaine and menthol. These findings, that pain disappears under the influence of anesthetics while the other senses remain and that it remains while the other senses disappear, spoke for specificity.

(5) Mapping the distribution of pain spots on the body with punctiform stimulation and graded esthesiometers.

(6) Liminal determinations of pain in different areas.

(7) The uniqueness of certain areas in yielding pain alone (*e.g.* the cornea of the eye) and the insensitivity of certain other areas (*e.g.* Kiesow's area, the mucous membrane of the cheek opposite the second lower molar).

The evidence for the sensory theory in hand in 1896 was impressive. It convinced many, but it did not place the theory in the realm of fact. There were still many unsolved problems. Outstanding among them were: (1) the definite identification of the pain-receptors; (2) the establishment of the peripheral pathways; (3) the localization of the cerebral centers; (4) the interpretation of the non-adaptability of pain; and (5) the explanation of its unique relationship with unpleasantness.

[68] Titchener, *Experimental Psychology: I. Qualitative Experiments: ii. Instructor's Manual*, 1901, 81.

Solutions of these problems were, to be sure, proposed in the '90s, but none were final. For example, Von Frey thought that the free nerve-endings were the organs of pain; but, his direct experimental evidence being meager, he relied upon circumstantial evidence. The free endings are the only nerve structures widely scattered throughout the body and existing in sufficient profusion in the skin to correlate with the numerous pain-spots localized upon the surface. Again, the non-adaptability of pain was explained biologically. Pain is deleterious; adaptation would be of anti-survival value, as organisms that become adapted to pain would, in the long run, not survive; hence pain is non-adaptable. A false explanation, to be sure, but one that satisfied the exigencies of the occasion.[64]

The sensory theory was widely accepted. When Sherrington[65] in Schäfer's *Textbook* and Thunberg[66] in Nagel's *Handbuch* wrote of pain as a specific sensation coördinate with the other cutaneous sensations, the theory was 'in,' and it was presented and taught in most textbooks of physiology and psychology as established fact. The other theories were eclipsed. The traditional pleasure-pain theory underwent a complete reorganization. After vain attempts to discover nerves of pleasure—Nichols, for example, searched for them throughout the body and finally decided that they must be in the organs of sex and alimentation, the two systems vitally concerned with the preservation of the species[67]—the theory was placed in a new perspective. Pain was removed as a term in the dichotomy, unpleasantness was substituted, and the dichotomy was changed to one of pleasantness-unpleasantness and then relegated to the affective sphere where it would no longer be of concern to those chiefly interested in sensation.

The intensive theory in its more general form was definitely disproved. Pain positively is not a common sense-quality that is aroused by excessive stimulation in all sensory nerves. The pains of dazzling lights, of shrill tones and intense noises, of pungent tastes and smells, of extremes of cold and heat, are mediated by the co-excitation of other nerves than those of the specific senses. So much we may accept as established fact.

Goldscheider's modification of the intensive theory was, however, un-

---

[64] Cf. also G. W. A. Luckey, Some recent studies of pain, this JOURNAL, 7, 1895, 108-123; W. von Tschisch, Der Schmerz, *Zsch. f. Psychol.*, 26, 1901, 14-32.

[65] C. S. Sherrington, Cutaneous sensations, in E. A. Schäfer's *Textbook of Physiology*, 2, 1900, 920-1001.

[66] Torsten Thunberg, Physiologie der Druck-, Temperatur- und Schmerzempfindungen, in W. Nagel's *Handbuch der Physiologie des Menschen*, 3, 1905, 647-733.

[67] Nichols, *op. cit.*, *Philos. Rev.*, 1, 1892, 414.

touched by these considerations. It was, nevertheless, at the textbook level at least, completely ignored; it was put in the discard along with the general form of the intensive theory with which it was incorrectly identified. At the frontiers of research, on the other hand, the theory was very much alive. There it took its place as a theory contending on equal footing with Von Frey's; and many have been the studies undertaken to decide between them!

Goldscheider partly prejudiced the issue by his adherence to the term *Gemeingefühl,* which is the context into which he set his theory. It is, nevertheless, clear from his treatment that he, like Von Frey, regards pain as a specific sensation subserved by a special class of nerves. From there on, however, their theories diverge. Goldscheider holds that pain is mediated by the tactile nerves and that it results from the summation of their excitations in the gray substance of the spinal cord. Von Frey, as we have seen, holds that pain is unique, that it is a separate modality subserved by discrete receptoral structures—the free nerve-endings. In writing of these theories, Kiesow says: "the differences between them goes so deep that a reconciliation is hardly conceivable . . . if one view is correct, the other must necessarily be false. Consequently, we cannot expect the conflict to end until one of the views attains general recognition and so flict to end until one of the views attains general recognition and so vanquishes the other."[68] But how to set the experimental stage so that the contending theories could meet? That question was long in answering.

As we have seen, pain, for biological reasons, was not supposed to show adaptation.[69] This point of view was held for many years. It was not until comparatively recent times (1919) that Straus and Uhlmann discovered in the Cornell laboratory that pain *did* show the phenomenon of adaptation.[70] One of the outstanding problems regarding pain, *i.e.* the explanation of its none-adaptability, was solved, and pain, as far as adaptation is concerned, stepped into line with the other cutaneous senses. All that is required for its adaptation is a relatively constant, unvarying stimulus—the very conditions necessary for adaptation in all the other departments of sense. When this result was obtained, the reason that pain had heretofore been regarded as non-adaptable immediately became clear. The stimulus-conditions in injuries, tooth-aches, and headaches are constantly varying.

[68] Friedrich Kiesow, The problem of the condition of arousal of the pure sensation of cutaneous pain, *J. Gen. Psychol.,* 1, 1928, 199-212.

[69] Titchener (A Textbook of Psychology, 1909, 154) says, in 1909, that "the pain sense does not appear to show the phenomenon of adaptation;" and Fröbes (*op. cit.,* 151) in 1923, writes, "der Schmerz zeigt keine Adaptation," although numerous experiments had by then been reported showing the contrary.

[70] H. H. Straus and F. R. Uhlmann, Adaptation of superficial pain, this JOURNAL, 30, 1919, 422-424.

Straus and Uhlmann were content to rest their study at that point, but the experiment, as it turned out, was destined to play a larger rôle. As an *experimentum crucis,* it led to a decision between Von Frey's and Goldscheider's theories. If Von Frey's sensory theory were correct, pain should, in the course of its adaptation, simply become progressively weaker and disappear without undergoing any qualitative change. If Goldscheider's theory, on the other hand, were correct, then, as pain becomes less intensive through adaptation, it should also undergo a qualitative change, *i.e.* it should change into pressure. Goldscheider,[71] accepting the logic of the adaptation-experiment, claimed that the excitation of the pain-spots arose and disappeared through subpainful experiences ("die Erregung der Schmerzpunkte unterschmerzlich beginnt und unterschmerzlich aufhört").[72] He found, in his studies of its adaptation, that pain not only became weaker under continuous stimulation but also that it underwent a qualitative change, *i.e.* before completely adapting and disappearing, it passed into a subpainful sensation (*unterschmerzliche Empfindung*) of contact (*Berührung*) or pressure (*Druck*).

In an experiment on pain[73] in 1922, Von Frey tested Goldscheider's subpainful sensations. In order to avoid the concomitant pressures that are aroused when an object is brought to bear upon the skin, Von Frey elicited pain by focusing the sun's rays upon the skin by means of a hand lens. Because he was able to obtain pure pressureless pains by this as well as by other means, he concluded that Goldscheider's summation theory was untenable.

In refutation, Goldscheider, in 1926, pointed out that Von Frey had used strong stimuli, and that the course of experience, under those circumstances, is telescoped and the subpainful pressures are obscured by the intense pains.[74] He laid down two necessary conditions for the subpainful pressures: first, the pain should not be intense; and secondly, it should be produced slowly, not quickly or suddenly. His conclusion was that "the doctrine of specific peripheral nerves of pain is unwarranted."[75] So the situation was back again to where it was before! Opinion set against opinion, and observation against observation.

In 1931, Wells and Hoisington reported a phenomenological study of pain-adaptation which, they thought, settled the controversy by reconciling the disparate points of view.[76] They found that the end-quality in the course of pain-adaptation

[71] Goldscheider, *Das Schmerzproblem,* 1920, 20, 71-73.
[72] *Op. cit.,* 20.
[73] Von Frey, Versuche über schmerzerregende Reize, *Zsch. f. Biol.,* 76, 1922, 1-24.
[74] Goldscheider, Beiträge zur Physiologie der Gemeingefühle, *Zsch. f. Sinnesphysiol.,* 57, 1926, 1-14.
[75] *Ibid.,* 14.

"is neither a weak pain (Von Frey) nor a pressure (Goldscheider), but a non-pressury 'bright contact'."[77] Upon the basis of this finding, they suggested that the difference between Von Frey and Goldscheider was principally terminological. Though recognizing that "interpretation and comparison of qualitative terms is always dangerous," they, nevertheless, ventured the opinion that Von Frey's "weak pains" and Goldscheider's "subpainful pressures," which those investigators obtained just before complete adaptation was reached, corresponded to their own "bright contacts." If such is the case—"if Von Frey's insistence that isolated pain spots give nothing but pain is . . . due to a confusion of 'weak pain' with non-painful 'bright contact'," and "if Goldscheider's contention that isolated pain spots give 'pressure' is the result of a failure to differentiate between true pressure and the non-pressury contact-like experience"—then, they conclude, "the controversy will be settled." Yes, settled by the elimination of Von Frey! For Von Frey to admit "that the final phase of pain subsidence is not really painful"—as Wells and Hoisington would have him do—would be for him to accept Goldscheider's theory. Wells and Hoisington reconcile the controversy by corroborating Goldscheider!

It was at this point that I undertook, in coöperation with several of my students, a series of studies on the problem. The first study (with Burns) was performed, following Goldscheider, with needle esthesiometers.[78] We found, in confirmation of Goldscheider, that pain was replaced by pressure in every instance before complete adaptation was secured. The normal and typical course of adaptation was "a gradual subsidence from a maximal pain through pressure to indifference."[79] At which point or points in the curve, however, was complete adaptation attained? When pain disappeared and pressure emerged? Or only when pressure disappeared? Or have we here two cases of complete adaptation: one of pain, and another of pressure? Our answers to these questions will depend upon our point of view, but it should be recalled that our stimuli—needle esthesiometers—were applied at superliminal pressure values; they were, therefore, capable of eliciting pressure as well as pain. The pressures reported after the subsidence of pain may very well be nothing more or less than the residual effect of our own stimuli. The only way of deciding the matter was to arouse pain with stimuli, like radiant heat, that do not give rise simultaneously to pressure.

We therefore devised, in the second study (with Stone),[80] a thermo-esthesiometer that yielded pain in conformity with the strict conditions laid down by Goldscheider for the observation of his subpainful pressures.

[77] E. F. Wells and L. B. Hoisington, Pain adaptation: A contribution to the Von Frey-Goldscheider controversy, *J. Gen. Psychol.*, 5, 1931, 352-366.

[77] *Ibid.*, 364.

[78] Maryland Burns and K. M. Dallenbach, The adaptation of cutaneous pain, this JOURNAL, 45, 1933, 111-117.

[79] *Ibid.*, 115.

[80] L. J. Stone and K. M. Dallenbach, Adaptation to the pain of radiant heat, *ibid.*, 46, 1934, 229-242.

Goldscheider said that there should be no telescoping of experiences and, as the typical reports of our observers showed, there was none. Warmth, the first quality elicited, passed slowly through intensive gradations to heat; this, in turn, passed through intensive gradations to weak pain which grew in turn in intensity, remained constant at a high degree for a short time, and then gradually declined, and disappeared into heat, which in turn subsided into warmth. Pressure was noticeable by its absence. This result might lend support to a new theory that pain is a summation of warmth; but we preferred the explanation that the accompaniments of pain, whether pressure or warmth, are functions of the type of stimuli employed. Since a pressure stimulus yields residual pressure on its way to complete adaptation, and a warmth stimulus residual warmth, we hazarded the opinion that a cold stimulus would yield residual cold—and thus set the problem for our third study.

This study (with Edes) yielded the expected results.[81] A dry-ice stimulator elicited pain without telescoping the experience. From slowly increasing degrees of cool and cold, pain emerged gradually, grew in intensity, remained at a high degree for a time, and then gradually disappeared, subsiding into decreasing degrees of cold and cool. At no time during the course of adaptation to the cold stimulus was pressure reported by our observers.

Did our observers, in the two studies with temperature, miss Goldscheider's pressures? Were those pressures so weak that they escaped notice? We think not. Our *O*s were highly trained in cutaneous observation; it is hardly possible that all of them would have failed all the time to report those qualities if they occurred, particularly as our experimental conditions fulfilled all the requirements that Goldscheider laid down and should, therefore, have been highly favorable for their observation.

All the pain-adaptation studies reveal, as residual effects, the qualities that are inherent in the stimuli. A pressure stimulus yields residual pressure; a warmth stimulus, residual warmth; and a cold stimulus, residual cold. Surely, these experiments justify us in concluding with Von Frey that the nerves of pain are separate and distinct from the tactile nerves, and that Goldscheider's summation theory is untenable. This conclusion is corroborated by experiments during the past decade in the field of neurophysiology, demonstrating that pain is subserved by specific nerve fibers.[82]

---

[81] Barbara Edes and K. M. Dallenbach, The adaptation of pain aroused by cold, *ibid.*, 48, 1936, 307-315.

[82] E. D. Adrian, *The Mechanism of Nervous Action*, 1932, 42-60.
H. S. Gasser and J. Erlanger, The rôle of fiber size in the establishment of a nerve block by pressure or cocaine, *Amer. J. Physiol.*, 88, 1929, 581-591.

Though the investigators disagree among themselves regarding details, they agree in general that separate nerve fibers subserve the function of pain.[83] Further support for this point of view is provided by the experiments in sensory chronaxy which demonstrate significant and reliable differences in the chronaxies and in the strength-duration curves of pressure and pain.[84] Thus we see that Nichols was right when he opined that the prejudice of science ran in favor of specificity.

Of the five outstanding problems at the turn of the century, I have touched upon two. One of these, as we have seen, has definitely been solved. The discovery that pain was adaptable removed the explanation of its non-adaptability from the list of problems. The other, the establishment of the pathways of pain, is well on the way to solution, if, indeed, it is not settled now. Of the three problems that I neglected, only one is nearer solution now than it was a third of a century ago. The identification of the pain-receptors is bound up with the establishment of the structural pathways, but the evidence for the free nerve-ending is only a little more convincing now than it was when Von Frey formulated his theory. The other two problems, the localization of the cerebral centers, and the explanation of pain's unique relationship with unpleasantness, are practically untouched. We still are, as Waterston recently observed, "entirely ignorant of any 'cortical area' associated with pain perception, nor do we know whether such an area exists or whether the thalamus alone . . . is sufficient for the pain sensation."[85]

The uneven advance since the turn of the century characterizes the history of the subject from the beginning.

P. Heinbecker and G. H. Bishop, The mechanism of painful sensations, *Proc. Ass. Res. Nerv. & Ment. Dis.*, 15, 1935, 226-238.

P. Heinbecker, G. H. Bishop, and J. O'Leary, Fibers in mixed nerves and their dorsal roots responsible for pain, *Proc. Soc. Exp. Biol. & Med.*, 29, 1932, 928-938; Analysis of sensation in terms of the nerve impulse, *Arch. Neurol. & Psychiat.*, 31, 1934, 34-53.

S. W. Ranson, Cutaneous sensory fibers and sensory conduction, *Arch. Neurol. & Psychiat.*, 26, 1913, 1122 f. 1 Cutaneous sensation, *Science*, 78, 1933, 397 f.

S. W. Ranson and H. K. Davenport, Sensory unmyelinated fibers in the spinal nerves, *Amer. J. Anat.*, 48, 1931, 331-353.

David Waterston, Observation on sensation: The sensory functions of the skin for touch and pain, *J. Physiol.*, 77, 1933, 251-257; On pain, *Lancet*, 224, 1933, 943-946; Pain and the mechanism of its production, *Brit. Med. J.*, (no. 3857), 1934, 1087-1089.

[83] The outstanding exceptions are Heinbecker, Bishop, and O'Leary (*opp. cit.*). These authors report results which agree with Goldscheider's theory.

[84] Cf. W. S. Neff and K. M. Dallenbach, The chronaxy of pressure and pain, this JOURNAL, 48, 1936, 632-637, and the references cited by them.

[85] Waterston, *op. cit.*, *Lancet*, 224, 1933, 943.

THE SPECTRAL COMPONENTS OF PSYCHOLOGICALLY UNIQUE RED

By FORREST L. DIMMICK and MARGARET R. HUBBARD, Hobart College

The determination of the spectral location of the seven 'psychologically unique' colors is important if the four-variable color equation is to become a valid means of description.[1] That is, the color equation acquires exact meaning only when the points of reference have been precisely standardized. In a previous paper we have discussed this question at length, and the standards for yellow, green, blue, black, white, and gray have been reported.[2] There remains, therefore, the definition of the remaining 'unique,' red. The majority of investigators who have studied the correlation of wave-length with hue have agreed that all spectral reds are yellowish. In the preceding experiment on the spectral location of the 'psychologically unique' colors we verified this observation.[3] Because the apparatus used in that experiment—a wavelength spectrometer—was suitable for work only with single wavelength bands, the present study was undertaken to determine experimentally the amount of 'unique' blue (476 mμ) that must be added to spectral red (635 mμ) to obtain the 'psychologically unique' red, *i.e.* a red which is neither bluish nor yellowish.

METHOD AND PROCEDURE

Apparatus. A colorimeter, designed and built in the Hobart laboratory for the purpose of obtaining well saturated colors by combining filtered light, proved satisfactory for our problem. (See Fig. 1) It consists of three optical systems A, B, C, all of which obtain light from a single source, S. The three systems converge and are focused upon a viewing screen coated with MgO. Light from the source falls directly, or is reflected by first-surface mirrors (M), upon three windows covered

* Accepted for publication May 3, 1939. Read before the Optical Society of America, February 24, 1939.

[1] F. L. Dimmick, A reinterpretation of the color pyramid, *Psych. Rev.*, 36, 1929, 83-90.

[2] F. L. Dimmick and Margaret R. Hubbard, An experimental determination of the spectral location of the psychologically unique yellow, green and blue, this JOURNAL, 52, 1939, 242-254.

Our attention has been called by Dr. Elsie Murray to an article not included in our previous review of the literature, in which an experimental determination of the yellow region is reported. J. G. Hailwood and H. E. Roaf (The sensation of yellow and anomalous trichromatism, *J. Physiol.* 91, 1937, 36-47) using a method of adjustment, determined the position in the spectrum of the points at which yellow becomes noticeably orangish and greenish. Their mid-yellow value is 583 mμ, which agrees with our results. Their range of variation is larger than ours, a fact which is undoubtedly a function of the difference in method.

[3] *Ibid.*

with fine grain ground glass (G), the image of which appears evenly illuminated on the MgO screen. The arrangement of lenses, (L), is such that the introduction of a variable diaphragm, (D), into the filter compartment, (F), between them, cuts down the illumination of the image on the viewing screen evenly over its whole area without altering the focus.

The observer views the central portion of the MgO screen through two reducing apertures (V_2V_3) by way of a first surface mirror (M_2) and a third aperture (V_1)

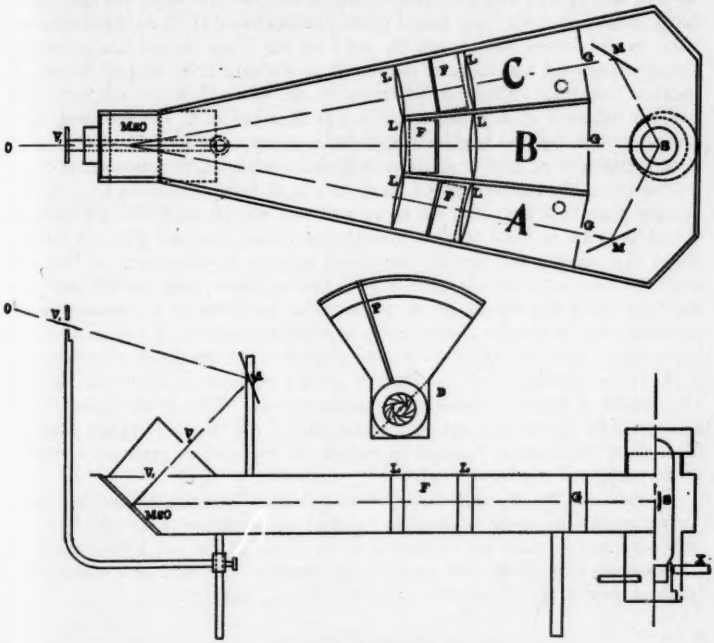


FIG. 1 HOBART COLORIMETER

which controls the line of vision. The field surrounding aperture V_2 is a middle gray whose reflectance is 0.200 which gives 5 foot lamberts brightness at an illumination of 25 f.c.;[4] those of V_1 and V_3 are black. The source of illumination (S) is a 300 w. T-10 projection bulb with a coiled coil filament rated at 115 v. and operated at 110 v. to give a color temperature of 3000° K. The lamp housing is cooled with a blast of air through an opening in the base (x).

The iris diaphragms (D) are calibrated in foot candles of illumination falling upon the MgO screen for every setting of the pointer (P). Since the apparent

[4] Dimmick and Hubbard, *op. cit.*, F. L. Dimmick and G. McMichael, A psychophysical determination of the limits of pure gray, this JOURNAL, 45, 1933, 314.

reflectance of the MgO screen is 1.00,[5] the brightness of the screen as viewed from O may be read in foot lamberts.

A Corning glass filter (No. 243, 2.19 mm.) which gives the equivalent of spectral red (635 mμ, purity 100%, brightness 9%) was placed in the filter compartment (F_a) of optical system A. Diaphragms B and C were set at zero and A was opened until there appeared on the viewing screen a well saturated, slightly yellowish red of medium brightness. The foot candle reading for this setting without the filter was 10.3. A blue color filter (Corning No. 554, 4.01 mm., 471 mμ, purity 100%, brightness 1.3%) was placed in filter compartment (F_c). As the diaphragm (D_c) in that system was opened, the color on the screen became less yellowish, passed through red, and appeared bluish red. At the same time, the total brightness increased somewhat. When the brightness change was sufficient to cut across the required judgment of hue, diaphragm A was decreased as C was increased, in a proportion that kept the brightness of the color constant. In the final series, sets of fixed apertures were used to give identical values in successive presentations.

Procedure. The colorimeter was set up in a dark room illuminated by a 500 w. Ivanhoe Northlight lamp, and the entire apparatus was screened from the observer except as much as could be seen through the viewing aperture (V_1). A shutter closed that opening save for the one second intervals of observation. A head-rest kept *O*'s head in a constant position. Observations were made entirely with the right eye which was maintained by the foregoing conditions in a constant state of moderate light adaptation. Before every experimental period, *O* remained in the experimental room for 5 min. to become adapted to the general illumination.

As in the preceding work, the *O*s were given a sufficient period of training with the method of limits to stabilize their judgments and to locate the region of the limens on the yellow side and on the blue side of red. The final values were derived from 100 series of 7 stimuli in each of the two regions, presented according to the method of constant stimuli.

Instructions. The *O*s were instructed as follows: "You will be shown a color. Judge whether the color is yellowish (bluish) or not-yellowish (not-bluish)." *O* was not asked to make any decision as to the purity of the red. Limens were determined for 5 of the *O*s who served in the previous experiment on the determinations of unique colors; namely, *C.D.*, *F.D.*, *F.*, *L.*, and *M.*

RESULTS

Table I shows the limits within which the red is neither yellowish nor bluish. Fig. 2 shows the location of the limiting values on the I.C.I. diagram. All *O*s agree so closely that the region is clearly delimited. Since there are no data on the number of J.N.D. within the limits of red, we have designated the mean of the two limits as 'red.' This gives 493.6c mμ for a representative red which is neither yellowish nor bluish. This value lies close to the one given by Exner; namely 494c mμ, and to the dominant

[5] D. B. Judd, Preparation and colorimetric properties of a magnesium oxide reflectance standard, *Nat. Bur. Stand.* LC-395, 1933, 3.

TABLE I

LIMITS OF THE REGION OF RED FOR ALL *Os*

Judged	*O*	Trichromatic coefficients λ	x	y	h	σ	Purity (%)	Brightness (ft. lam.)
Yellow in red	C.D.	493.4 c	.631	.253	(3.34)	(.212)	86	.95
	F.D.	492.7 c	.662	.264	(2.46)	(.287)	87	.94
	F.	493.1 c	.645	.258	(2.61)	(.271)	86	.95
	L.	493.5 c	.627	.250	(2.34)	(.302)	86	.97
	M.	492.8 c	.657	.262	(2.32)	(.305)	87	.94
	Av.	493.1 c ± .3						
Blue in red	C.D.	495.7 c	.550	.222	(2.86)	(.247)	83	.96
	F.D.	494.1 c	.600	.240	(1.56)	(.453)	85	.96
	F.	493.7 c	.618	.248	(1.65)	(.429)	85	.96
	L.	494.8 c	.577	.234	(4.31)	(.164)	83	.99
	M.	493.7 c	.619	.247	(1.24)	(.570)	86	.96
	Av.	494.4 c ± .7						

Median Value for Red 493.6 c

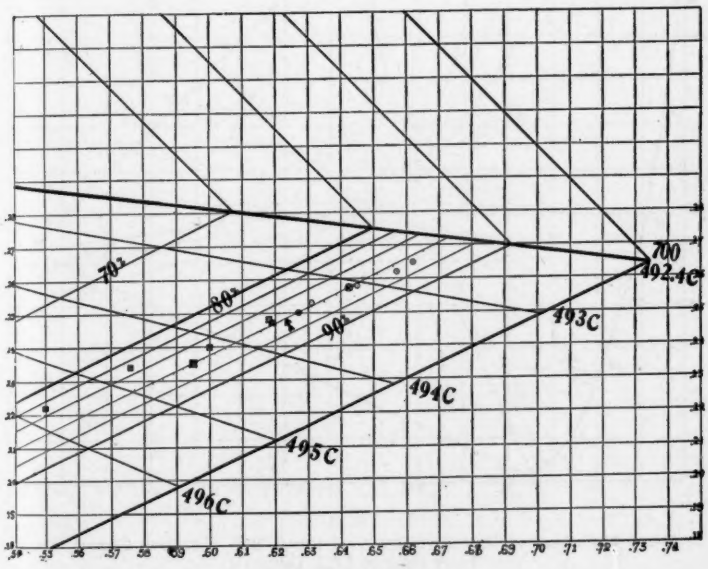


FIG. 2

FIG. 2 LOCATION OF THE LIMITS OF RED ON THE I.C.I. DIAGRAM
Small squares = 'just-not-blue' limits for all *Os*. Large square = Average 'not-blue' limit. Small circles = 'just-not-yellow' limits for all *Os*. Large Circle = Average 'not-yellow' limit. Arrow = Median value.

wave length of the Munsell sample 10-RP 5/10, though the latter is somewhat less saturated.

With the determination of 'unique' red as the complement of 493.6 mμ,

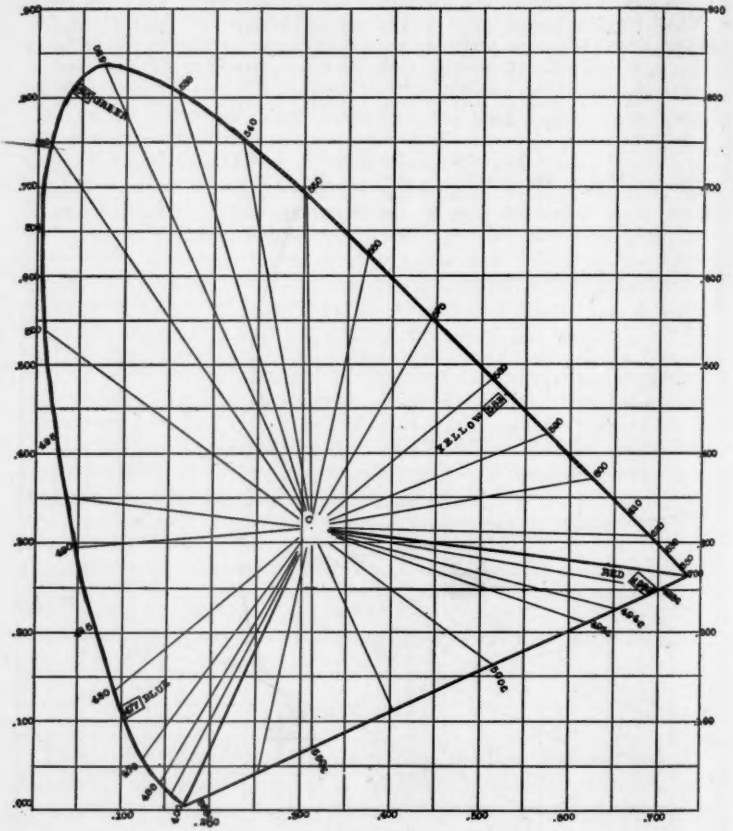


FIG. 3

FIG. 3 LOCATION OF PSYCHOLOGICALLY UNIQUE RED, YELLOW, GREEN, AND BLUE ON THE I.C.I. DIAGRAM

we now have the physical specifications of the seven psychologically unique colors (see Fig. 3). In terms of these, it is possible to put into quantitative form the four-variable color equation, x (red or green) ± y (yellow or blue) ± z (black or white) + G (Gray) = C.[6] Since the specifications

[6] Dimmick, *op. cit.*

are given in trichromatic coefficients, transformations can be made from one notation to the other and comparisons can be made with other systems of notation which have been or can be reduced to trichromatic coefficients. Full quantification of the psychological color equation now requires the determination of the J.N.D. constants relative to the seven points of reference.

Summary

The amount of 'unique' blue (476 mμ) that must be added to spectral red (635 mμ) in order to produce a red which is neither yellowish nor bluish was determined. The just-not-yellow and just-not-blue limits of red were obtained with 5 *Os*, and a representative red was calculated as the mean between these two limits, yielding the value 493.6c mμ. The seven points of reference in the four-variable color equation have now been experimentally established.

CHANGES IN TENSION-PATTERN AND TOTAL ENERGY EXPENDITURE DURING ADAPTATION TO 'DISTRACTING' STIMULI

By G. L. FREEMAN, Northwestern University

The problem for this investigation was suggested by two seemingly incompatible results on the energy costs of mental work. Morgan,[1] testing the effect of distraction upon reaction-time, reported increased performance and increased muscular tensions. Davis,[2] however, showed that repetition of noise-stimulation produced a fairly rapid decrease in the magnitude of the action-potentials recorded from the right forearm. Harmon,[3] in a preliminary examination of the same question, found increased metabolism per unit of work done when noise was first introduced, but noted that in a relatively short time the worker tended to return to the metabolic rate characteristic of work under conditions of quiet. Commenting on several possible explanations of these results, Poffenberger says "one might accept the increased tensions noted by Morgan as authentic and interpret the reduced tensions demonstrated by Davis as a *shift of tensions* to new muscles with a reduction of tension only in those muscles on which measurements were at the time being made. The decreasing metabolic cost as adaptation to the distraction progressed would indicate merely that the later tension patterns functioned more economically than the earlier ones."[4]

Evidence for such shifts in tension pattern were found in experiments by the present writer both during the acquisition of a new habit and under changes in instruction.[5] In the present report, we have applied our method of recording muscle tensions from the four limbs to the problem of adaptation to distraction and have added measurements of total energy expenditure in terms of oxygen consumption. The results obtained may be regarded as a preliminary analysis of the compensatory changes in tension pattern

* Accepted for publication September 6, 1938. Studies in Muscular Tension; Series 2, No. 9.

[1] J. J. B. Morgan, Overcoming distraction and other resistances, *Arch. Psychol.*, 5, 1916 (no. 35), 1-84.

[2] R. C. Davis, The muscular tension reflex and two of its modifying conditions, *Indiana Univ. Pub., Sci. Ser.*, no. 3, 1935, 1-23.

[3] F. L. Harmon, The effect of noise upon certain psychological and physiological processes, *Arch. Psychol.*, 23, 1937 (no. 147), 1-81.

[4] A. T. Poffenberger, Some unsolved problems in human adjustment, *Science*, 87, 1938, 124-129.

[5] G. L. Freeman, The spread of neuro-muscular activity during mental work, *J. Gen. Psychol.*, 5, 1931, 479-492.

which act to maintain the economy of adjustment even under conditions of distraction.

The Experiments

Two adults (*F* and *G*) performed addition problems under conditions of quiet and noise stimulation of approximately 50 db. on 12 consecutive days. Work periods of 20 min. were divided equally between the two conditions, and were preceded by 40 min. of rest. Total oxygen consumption and action potentials from the four limbs were recorded during the last 10 min. of rest and also during the mental work. All experiments were done in the morning with the *Ss* in the post-adsorptive state (without food for 12 hr.). While the conditions of quiet or noise stimulation were alternated in a random order so that records obtained in the basal state of rest would not reflect specific anticipation of noise or quiet, half of the time noise stimulation preceded quiet work and half the time followed it. (In preliminary series it was found that if *S* expected noise stimulation to precede quiet work, his neuro-muscular set or preparation differed from that which accompanied expectation of quiet working conditions.) During the experiments *S* lay in a semi-prone position on a cot. The addition problems (10 three-place numbers) were printed on cards and *S* was instructed to put his sums upon a long strip of paper attached to his writing board. The Jones motor-basal apparatus[6] was used for the metabolic measurements and the muscle action potentials were obtained from small (6 mm.) gold-plated disks cemented to the dorsal surface of the upper arms and legs. A common 'indifferent' lead was attached to the abdominal region, and a commuting arrangement permitted serial recording from the four areas at the rate of 2 per sec. The common amplifier used with these leads was built after the model described by Davis[7] and was connected to a crystal oscilloscope (type S_1) with optical level arranged for photographic recording. The physiological measurements made during each condition of rest, 'quiet' work and 'noise' work were, therefore, 2 separate determinations of the metabolic rate and 10 momentary records of action potentials from each of the four muscle groups.

Results

In Table I we present the addition scores under conditions of quiet and noise distraction. As in the study by Harmon,[8] we find only a temporary reduction in work output upon the introduction of noise distraction. The percentages of correct solutions were practically identical in the two situations, the average for quiet being 80.4% and for noise 78.8%. The critical ratios of the experimental differences were low for all days and decreased progressively in magnitude with adaptation of the subject.

Since the *Ss* attempted to work at top speed in both conditions and were unaware of their scores, it is possible that compensatory effort was utilized during distraction. Records of oxygen consumption, given in Table II, do

[6] The Jones motor-basal apparatus, *J. Amer. Med. Ass.*, 103, 1934, 1620.
[7] R. C. Davis, *op. cit.*
[8] F. L. Harmon, *op. cit.*

indicate a considerable increase in total energy expenditure, especially during the first few days of work under noise stimulation. Its gradual reduction to somewhere near the level originally established for quiet-work is shown by the Q-N scores. All metabolic measurement in the table has been reduced to the calories consumed per min. and the values given for conditions of basal rest, quiet-work, and noise-work are the average of two tests made daily under each condition. Records taken during the basal rest have

TABLE I

NUMBER OF CORRECT ADDITIONS DURING THE DIFFERENT 10-MIN. PERIODS

Day	Quiet			Noise		
	F	*G*	Av.	*F*	*G*	Av.
1	30	28	29.0	26	29	27.5
2	30	35	32.5	28	27	27.5
3	37	54	45.5	34	45	39.5
4	40	38	39.0	43	41	42.0
5	45	40	42.5	41	36	38.5
6	40	50	45.0	38	48	43.0
7	41	42	41.5	40	45	42.5
8	43	48	45.5	42	46	44.0
9	40	56	48.0	44	55	49.5
10	41	48	44.5	41	52	46.5
11	38	55	46.5	36	55	45.5
12	40	47	43.5	42	49	45.5

a reliability coefficient of 91.4± .12, while records taken under conditions of quiet-work and noise-work have a test-retest correlation of 70.5 ± .15 and 81.0 ± .13.

The slight but consistent increase in metabolic activity which accompanies mental work is in accordance with the expectation derived from studies by Benedict and Benedict,[9] Harmon[10] and others.[11] Such increments are insignificant when compared with the demands of physical work, and it is probable that no surplus store of fats or carbohydrates is drawn upon and that little oxygen debt is incurred.

The introduction of noise-distraction in such a situation causes at first a marked increase in metabolic activity. With continued stimulation, however, this effect is obscured by the readjustment of the individual to the new working condition, and distraction *per se* has little far reaching effect so far as general bodily economy is concerned. For an analysis of internal reorganizations which occur in the adaptive process we turn to action potential records from various muscle groups.

It will be recalled that we obtained action potential records from the

[9] F. G. Benedict and C. G. Benedict, The energy requirements of intense mental work, *Proc. Nat. Acad. Sci.*, 16, 1930, 438-443.

[10] F. L. Harmon, *op. cit.*

[11] F. C. Becker and O. Ollson, Metabolism during mental work. *Skand. Arch. f. Physiol.*, 31, 1914, 80-197.

upper arm and leg muscles at 2-sec. intervals during the conditions of basal rest, quiet-work, and work under noise distraction. In Table III these records are treated as if they were simultaneous and are scored in terms of the total deflections per sec. according to the method advocated by Davis.[12]

TABLE II

THE EFFECT OF BASAL REST, QUIET-WORK, AND WORK UNDER NOISE DISTRACTION UPON TOTAL EXPENDITURE OF ENERGY

(Calories consumed per min.)

S	Condition	Days of tests											
		1	2	3	4	5	6	7	8	9	10	11	12
F	Basal	1.61	1.87	1.64	1.50	2.13	1.92	1.54	1.61	1.87	1.82	1.63	1.70
	Quiet	1.86	1.91	1.64	1.62	2.10	2.00	1.63	1.68	1.82	1.80	1.78	1.73
	Noise	2.24	2.10	2.18	2.25	2.20	2.16	1.72	1.52	2.01	1.78	1.93	1.73
	Diff.$_{N-Q}$	.37	.18	.54	.63	.10	.16	.08	−.17	−.19	−.02	.14	.01
S	Basal	1.78	1.60	1.60	1.82	1.77	1.67	1.62	1.78	1.60	1.75	1.82	1.82
	Quiet	1.72	1.71	1.69	1.80	1.79	1.84	1.75	1.72	1.65	1.52	1.83	1.76
	Noise	2.43	2.53	1.82	2.41	1.92	1.83	2.24	2.06	1.65	1.73	2.02	1.80
	Diff.$_{N-Q}$	.71	.82	.12	.61	.13	−.01	.45	.33	−.00	.10	.18	.04
Av.	Basal	1.69	1.74	1.62	1.66	1.95	1.80	1.58	1.70	1.74	1.78	1.72	1.76
	Quiet	1.79	1.81	1.67	1.71	1.95	1.92	1.69	1.74	1.74	1.71	1.31	1.78
	Noise	2.34	2.32	1.99	2.33	2.06	1.99	1.98	1.79	1.83	1.75	1.97	1.77
	Diff.$_{N-Q}$	.54	.50	.33	.62	.11	.08	.26	.08	.09	.04	.16	.02

TABLE III

MAGNITUDE OF THE ACTION POTENTIALS IN THE FOUR LIMBS ON SUCCEEDING DAYS, DURING BASAL REST (R), QUIET-WORK (Q), AND WORK UNDER NOISE DISTRACTION (N)

(Total deflections per sec., average of 2 Ss)

Days	Right arm			Left arm			Right leg			Left leg			Average		
	R	Q	N	R	Q	N	R	Q	N	R	Q	N	R	Q	N
1	120	361	525	97	225	742	102	200	904	165	215	835	121	250	751
2	85	228	444	109	287	801	150	225	504	154	291	520	124	258	569
3	125	312	518	150	265	628	85	254	475	152	247	515	128	269	534
4	110	276	432	67	192	485	107	194	312	175	163	388	115	206	404
5	254	295	449	285	217	527	318	225	350	280	297	291	284	258	404
6	148	224	421	95	210	433	205	212	288	227	238	245	169	221	347
7	72	200	344	65	172	405	188	175	183	105	189	192	107	184	281
8	103	287	370	97	154	375	150	142	180	174	170	178	131	188	276
9	125	195	362	136	188	291	207	217	231	159	143	140	157	186	256
10	96	193	394	114	177	215	227	225	212	248	220	187	171	204	252
11	50	225	345	87	103	200	135	243	154	127	151	163	100	180	215
12	200	219	337	182	185	282	212	177	170	181	188	175	194	192	241
Av.	124	251	413	124	198	449	174	207	330	179	209	320	150	216	377

Daily averages are given under each condition. Mean deviations of individual averages ranged from 84 to 324. In spite of this variability, the data still appear to be highly consistent for a given *S* in all three situations; that is, there was relatively less electro-myographically recorded activity shown during basal rest than in quiet-work and noise-work conditions.

[12] R. C. Davis, *op. cit.*

Table III indicates that while the right arm muscles were usually most active in the condition of quiet-work, their activity is exceeded in the condition of noise-work by other muscles groups, especially during the first few introductions of the distracting stimuli. This shift in the pattern of muscle tensions is not open to a 'fatigue' interpretation, for in half of the cases the noise condition actually preceded the quiet condition without producing any marked change in the general trend. What is apparently exhibited is an habituation effect to a new and novel situation, the condition of quiet

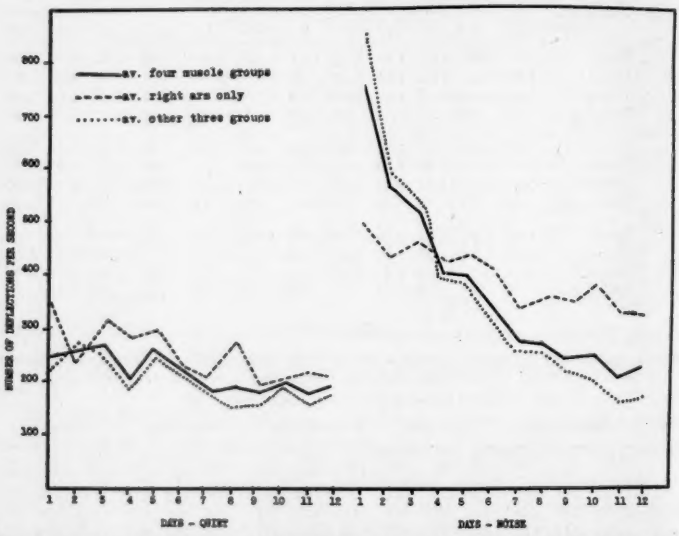


FIG. 1. AVERAGE MAGNITUDE OF ACTION POTENTIALS ON SUCCESSIVE DAYS UNDER QUIET WORKING CONDITIONS AND UNDER NOISE DISTRACTION

work presumptively occasioning a more usual or normal performance for the subject.

Fig. 1 shows clearly the redistribution of the tension pattern upon the introduction of noise and also indicates the progressive change which this pattern undergoes as *S* adapts to the distraction. The solid line represents magnitude of action potential deflections when the activity of all four muscle groups is averaged; it shows little change from day to day during quiet work while for noise it falls rapidly from an initial peak so that toward the end of the tests total activity is on the level established for quiet work. As might be expected, these values correlate positively with the determinations of total energy expenditure in the same situations (0.51 for quiet-work and 0.74 for noise-work).

Because both *Ss* used the right arm in totaling sums and because there was a tendency in all conditions for muscular activity to be more concentrated there than in other bodily parts, we have separated these action potential scores from the totals and indicated their comparative magnitudes by means of a broken line. The dotted line, also appearing on the same graph, is the average magnitude of action potential activity occurring correlatively in the other three limbs. Taken *in toto,* the graph suggests that in working adaptation to noise, activity in the focal part of the supporting tension pattern (right arm) is slightly increased (over the quiet work control) and remains so during the period of test. The more obvious change under distraction is in what might be called the non-focal parts of the tension pattern. Activity in the other limbs, while normally less than that of the right arm, is markedly increased at the first introduction of distraction. An interpretation of the parallelism between progressive changes in these parts and in total energy expenditure is that such activities are compensatory and tend to drop back to a minimum as the focal part of the tension pattern reëstablishes a more economical function.

Discussion

The results obtained in this study suggest that the initial effect of distracting stimuli is to temporarily unbalance a bodily economy in which, for habituated acts at least, supporting processes operate at an optimal level. The added effort which *S* applies to compensate for the distraction expresses itself not only by a rise in total energy expenditure, but also by a shift in the pattern of supporting processes. There is a marked spread of neuromuscular activity to parts but remotely connected with the task at hand, and the original focus of the pattern tends to be lost. Under repetition of stimulating conditions, however, this 'irradiation' effect becomes less prominent and the more economical concentration of supporting processes in limited areas begins to reassert itself. Such redevelopment of a focalized response-pattern may involve a slightly greater degree of supporting tensions, but these changes apparently are not extensive enough to have much effect on the economy of total energy expenditure.

The fact that we obtained records showing a concentration of neuromuscular activity in the right forearm during addition under quiet work may well be an artifact of our conditions;[13] that is, utilization of some other effector part for setting down the sums might have developed a greater

[13] R. C. Davis (*op. cit.*) obtained records from the right forearm showing marked reduction under noise distraction. However, it should be recalled that Davis was not dealing with the problem of working distractions; that is, noise was introduced upon a condition of rest, in which differential tension patterns might not have been as prominent as in the present experiment.

degree of activity in that part. Even were this supposition true, it would seem not to weaken the principle of redistribution in tension-load under distracting stimulation.

In an earlier study it was found that habituation of an act (though not learned under noise conditions) was accompanied by a reduction in tension in remotely related muscle groups.[14] This work has subsequently been confirmed and developed by others.[15] The theory and practice of differential relaxation for various tasks recognizes similar principles.[16] One research in which simultaneous records were taken from four muscle groups showed that habituation produced a concentration of tensions in muscles primarily involved in the task.[17] It would seem that the present experiment is an extension of this observation to the problem of work under distracting conditions, the supporting processes at first outbalance the distraction by intensive and extensive increments. During protracted work under these conditions, however, a more economical pattern is reëstablished. Work output during the entire readjustment process may remain fairly constant, and according to present indications at least, the net effect upon total energy expenditure may also be minimal.

While these results are suggestive, it must be obvious that they leave many related problems untouched and unsolved. Especially we would call attention to the fact that interpretation of the *shift in tensions* under habituation is not yet understood in terms of neural dynamics. Also, it is not known how effectively such shifts would operate either for heavier working loads or under more intense distraction. The *Ss* in this experiment were perhaps not as seriously disturbed by noise as many others, or they may have possessed more stable or efficient neuro-muscular systems. To hypothecate such limitations is at present both unwarranted and unwise; but that these may exist even in young children is suggested in studies by Johnson,[18] Duffy[19] and others.[20]

---

[14] G. L. Freeman, *op. cit.*

[15] J. Stroud, The rôle of muscle tensions in stylus maze learning, *J. Exper. Psychol.*, 14, 1931, 606-631; M. S. Clites, Certain somatic activities in relation to successful and unsuccessful problem solving, *ibid.*, 18, 1935, 708-724. E. Ghiselli, Changes in neuro-muscular tension accompanying the performance of a learning problem involving constant choice time, *ibid.*, 19, 1936, 91-99.

[16] E. Jacobson, Electrophysiology of mental activities, this JOURNAL, 44, 1932, 677-794.

[17] G. L. Freeman, *op. cit.*

[18] B. Johnson, Changes in muscular tension in coördinated hand movements, *J. Exper. Psychol.*, 11, 1928, 334.

[19] E. Duffy, The measurement of muscular tension as a technique for the study of emotional tendencies, this JOURNAL, 44, 1932, 146-162.

[20] G. L. Freeman, and E. T. Katzoff, Muscular tension and irritability, *ibid.*, 44, 1932, 789-792; M. A. Wenger, Some relationships between muscular processes and personality, *Child Dev.*, 9, 1938, 261-276.

PERSISTENT ALPHABETICAL SYNESTHESIS

By H. L. HOLLINGWORTH, Columbia University, and VIRGINIA WEISCHER, Duke University

An eight-year-old child was heard to remark, upon first hearing the name of an unfamiliar person, "That name is brown, isn't it?" Further inquiry revealed that all names were colored in her experience and that she supposed that this was the experience of others also. A series of verbal stimuli brought out the promptness of her color responses, and these were recorded. A year later the same series was repeated, and the responses were so similar to the earlier ones that it appeared desirable to make systematic record of them, and to make some study of their consistency and permanence.

The initial observations were made 12 yr. ago (1926) when *S* (the junior author) was 8 yr. old. The initial stimuli were presented again at the ages of 9 and 10 yr., without previous notice. Thereafter nothing was said of the matter for a period of 10 yr. In December 1938, *S* was 20 yr. old, a college senior, and an opportunity occurred to repeat the initial experiment, and also certain extensions that had been added at the trial 10 yr. previous.

Since the main interest of the earlier repetitions of the stimulus-series lay in the question of the persistence of the color phenomena, no very systematic exploration was outlined. In the light of the later findings it would have been better had a more thorough going alphabetical analysis been made in the beginning. But two sets of responses, 10 yr. apart, subsequent to two earlier annual tests, seem to bring out the essential features of the phenomena, even if they do not suggest a wholly adequate explanation of them.

The available data are as follows. Stimuli presented 4 times, over a period of 12 yr., from childhood to maturity, were the nine digits, three two-number combinations, the days of the week, seven proper nouns, an abbreviation, and inquiries concerning the sound of two musical instruments. The color responses to these are shown in four columns of Table I.

In 1928, additional stimuli were included in the series, and all these were repeated at the final test, 10 yr. later. These included nine words whose meanings were unknown to the child (Table II), eight pairs of words sounding alike but having different meanings (Table III), and fifteen reversed questions, in which the color was named by the experimenter and *S* asked to state what names have that color (Table IV).

In the 1938 examination it appeared that the key to the results was to be found in the fact that each digit and each letter of the alphabet had from the beginning (and for some time previous to the examination at 8 yr. of age) had a specific color, the colors of words being those of the dominant or initial letters. At this time the color of each letter of the alphabet, in random order, was asked for, and a few days later this alphabet series was repeated (Table V).

In addition *S* was asked various questions about the nature of the experience, and in 1938 wrote out a brief account of its character. In these accounts it appeared that

* Accepted for publication January 9, 1939.

TABLE I

SYNESTHETIC RESPONSES FROM CHILDHOOD TO ADULT LIFE

Stimulus-word, letter, or number	Color responses on following dates			
	Nov. 1926	Nov. 1927	Nov. 1928	Dec. 1938
One	white	white	white	white
Two		white	yellow	white and yellow
Three	yellow	dark yellow	brown	yellow
Four	pink	pink	pink	pink
Five	blue	blue	blue	blue and yellow
Six		black	black	black
Seven	yellowish brown	light olive	light olive green	green
Eight	reddish brown	brown	reddish brown	brown
Nine	dark green	now black often changes	black	black and purple
Monday	blue and white	blue and brown	blue	blue and white
Tuesday	yellow and black	yellow and black	yellow	yellow
Wednesday	gray	gray blue	gray and blue	bluish gray
Thursday	yellow and brown	yellow	yellow	yellow and blue
Friday	gray	most gray some blue	blue-gray	purple
Saturday	gray	gray	gray	black, I guess
France	yellow or purple	yellow and purple	purple	purple
Greece	green	green and gray	gray	green, then whitish
New York	white and gray	gray and white	gray and blue	gray and dark
London	yellow and green	yellow	yellow	yellow and white
Montrose	Mont—brown rose—pink	Mont—brown rose—pink	Mont—brown rose—pink	brown
Virginia	pink	pink	pink	pink
Hollingworth	Holl—purple ing—yellow worth—white and black	purple, yellow orange gray	purple yellow and pink gray	"blue and black—the worth is light yellow or light"
Mrs.	M—lavender r—green s—red	brown green red	"red and yellow"	brown brown red
"the sound of a violin"	purple	purple	purple	"green, but can't be sure. The word is brown."
	(Says "Maybe because 'violin' and 'violet' are similar.")			
"the sound of a piano"	red or green or brown	high—yellow, pink low—brown, black or dark red	"None"	white
47	4 pink 7 green	pink yellow green	pink olive green	pink green
86	8 red 6 gray	brown black	reddish brown black	brown black
35	3 yellow 5 blue	light tan blue	yellow blue	yellow blue

the precise alphabetical character of the synesthesis was not recognized by *S*. The colors just "came," or they did not. A discussion of the results follows the tabulation of the responses.

The most striking things about these results are their remarkable consistency and systematic character, and their persistence from early childhood well into adult life

(from at least as early as 6 to later than 20 yr). The description of the experience in itself also has some interest.

When first observed, *S* was 8 yr. old, with an I.Q. of about 150. She had been from infancy interested in colors, and her mother reports that colors were the first things she knew the names of. She was fond of working with colors, in drawings and decorative designs, and had a good appreciation of color effects.

TABLE II

SYNESTHETIC RESPONSES, 10 YEARS APART, TO STIMULUS-WORDS WHOSE MEANINGS WERE UNKNOWN

Stimulus-words	Nov. 1928	Dec. 1938
synapse	red and green	*red and yellow
Fechner	purple and green	pink and yellow
psyche	red	*pink and white
glossopharyngeal	orange and purple	yellow and last part bluish gray
neurilemma	green and yellow	green and yellow and gray
Babinsky	brown	brown
cortex	white	*white and brown
aorta	blue and yellow	blue and white
Kegan	purple	green and orange

* Meaning of the stimulus-word now known.

TABLE III

SYNESTHETIC RESPONSES, 10 YEARS APART TO STIMULUS-WORDS SOUNDING ALIKE BUT WHOSE MEANINGS ARE DIFFERENT

Stimulus-words	Nov. 1928	Dec. 1938
tea	yellow	yellow
"t"	yellow	yellow
sea	red and white	red
see	orange	red and black
fast (swift)	blue and black	lavender
fast (not eat)	blue and black	lavender
post (of fence)	pink	pink and white
post (a letter)	pink	pink and white
blue	gray	blue and gray and white
blew	blue and yellow	blue, gray and white
sew	red and brown	red and black
so	red and white	red and white
chews	white and yellow	reddish and light
choose	white and yellow	reddish and light
Greece	gray	"Don't know,—they keep
grease	gray	changing."

She sees colors only in connection with letters, words, and numbers. The symbols themselves appear, apparently "in her head." They are not as if printed on a page, but there is a vague light background, and the symbols appear in color or in a spot of color. At the age of 20 yr. she says:

"As I hear the word it immediately seems to be printed in front of my eyes—and yet inside my head. Sometimes a word will appear to be all one color and sometimes it will have several colors."

"Colors do not appear when I read silently or aloud nor when others read or speak, usually. I never notice colors in words unless I *think of* a particular word by itself. In conversation an unfamiliar word sometimes occurs to me as having color because I know nothing else about it."

"Usually the first letters and the letters which have the most prominent sounds stand out from the rest and seem to have much more color. Often a great many of the letters in a word may be grayish or nondescript and a few of the important letters will give their color to the whole word."

TABLE IV

SYNESTHETIC RESPONSES, 10 YEARS APART, TO SPECIFIC COLORS

Stimulus—situation	Nov. 1928	Dec. 1938
What day is yellow?	Tuesday and Thursday	Tuesday
What day is gray?	Saturday ("Fri" would be gray but "day" not)	Friday
What day is gray-blue?	Wednesday	Wednesday and Monday
What day is red and yellow?	Sunday	Sunday
What number is blue?	5	5
What number is brown?	none	8
What number is white?	1	1
What number is black?	6 and 9	6 and 9
What number is yellow?	2	2, 3, 13, 30
What number is pink?	4	4
What cities are gray?	Athens and New York	New York
What cities are yellow?	Berlin	London
What cities are green?	Rio de Janeiro	"don't think of any"
What instruments sound purple?	violin	"violin might"
What instruments sound brown?	"none I think of"	drum

TABLE V

SYNESTHETIC RESPONSES, A FEW DAYS APART, TO THE LETTERS OF THE ALPHABET GIVEN IN HAPHAZARD ORDER

Letter	1st trial	2nd trial	Letter	1st trial	2nd trial
A	black and blue	black	N	green	green
B	brown	brown	O	white	white
C	white	white	P	pink	pink
D	green and brown	green and brown	Q	light gray	gray
E	black	black	R	green	green
F	bluish black	purplish gray	S	red	red
G	orange	orange	T	yellow	yellow
H	black	bluish gray	U	gray	gray
I	white	white	V	pink	pink
J	yellow	yellow	W	blue	grayish white
K	purple	gray	X	brown	brownish
L	yellow	yellow	Y	white	white
M	brown	brown	Z	gray	gray

Voices, musical tones, tastes, odors, and words in foreign languages do not ordinarily have color. Concrete objects have their own colors but their names, including proper names, have distinctive colors, not the colors of the objects of which they are the names. Color names are not exclusively the colors which they name, though mainly so. The color of 'blue' is "grayish or whitish and blue." But 'purple' is "purple."

Detailed examination of the tabulated responses brings out a number of interest-

ing things. Over the period of 12 yr. the digits maintain their characteristic colors, with only trivial variations. Even the variations are consistent—thus 7 is yellowish brown, light olive, light olive green, green. The variation is apparently in the momentary choice of names, rather than in the color tone itself.

The days of the week also have their consistent colors from childhood to adult life; mainly the colors of their initial capital letters, as may be seen by referring to the alphabet list. As for other words, "France" is mainly purple (the color of F), and "Virginia" (her name) is always pink (the color of V). She says she always wished the color of her own name to be blue but "It always stayed pink." It is not so easy to see where the greenish-gray of "Greece" comes from (perhaps from the R, C and EE, green, white, black) but the initial letter (G, orange) does not count here.

In some words each syllable has its color; in the abbreviation "Mrs." each letter is separately colored. The two-place numbers are invariably two-colored, the colors of their respective digits. Even words she cannot precisely spell, and whose meaning she does not know, have characteristic color. Thus, at 10 yr. of age, "neurilemma" is reported as "green and yellow" (N—green, L—yellow). "Babinsky" is of course brown, from its capital B.

Words which sound alike but have different meanings, do not vary in color. This might be due either to the identity of sound or to the identity of initial letter. But it is apparently not phonetic elements that have color for this *S*, it is instead the letters of the alphabet and the single digits. Nevertheless it is of some interest that all the vowels are achromatic—white, gray or black (including also the Y which is often a vowel sound). It is the consonants that have the spectral colors.

Explanatory conjectures readily occur, but exceptions are easily found to most generalizations that have been tried out. Thus it might be surmised that the associations are mediated by the letters of the color names. It is true, for example, that B is *b*rown; but so also are D,M,X. W*h*ite is the color of I and Y; but also of C and O. It might seem "natural" for L to be ye*ll*ow and for P to be *p*ink. But J and T are also yellow and 4 and V are pink too. The name of a color does not itself have that color exclusively. Hence this theory falls short.

The usual sensory and cortical theories that have been applied to the miscellaneous synestheses do not fit well with the fact that in this case sounds other than letters and numbers lack color.

Did *S*, as an infant, learn the letters and numbers from blocks or pages in which each had its fixed color, and in which all the vowels were achromatic? This theory has a reasonable ring, but neither *S* nor her mother has any recollection of any such playthings, and none are to be found in the collection of things she has preserved from earliest childhood. She is said to have learned to read at a very early age, by memorizing stories read to her from the ordinary printed page and then tracing these stories on the page, thus first learning the words, rather than beginning with the alphabet.

Nevertheless some such explanation fits well with the fact that the colored hearing in this case appears to go back fundamentally to the characteristics not of sounds as heard but of letters as "thought of" or "imaged." Such a genesis would articulate well with the most reasonable explanations offered for more casual and transient synestheses.

Summary. An eight-year-old girl reported that all words had their characteristic colors. She had apparently had this experience for at least two years or so. A list of stimulus-words was presented and the color responses recorded. One year and two years after the same list was presented. After a lapse of ten years from the last annual test, twelve years after the first examination, the observations were repeated, with a duplication also of certain new stimulus-words that had been added ten years before. The reports were surprisingly consistent and systematic. The key to the phenomena was found to lie in the fact that each digit and each letter of the alphabet had a special color, and these elements, or prominent or initial elements, gave the general color tone to words and two-place numbers. Unsuccessful attempts were made to trace the genesis of this alphabet synesthesis. *S* was not clearly aware of the very systematic character of the associations, which have persisted well into adult life.

POST QUANTIFICATION OF INTROSPECTIVE DATA

By Douglas Fryer, New York University

Descriptions of conscious activity during short tests of simple performance have been rated upon a scale of degrees of motivation and correlated with changes in performance from test to test as an indicator of the validity of this method of post quantification of introspective data.

Instructions. The general description of conscious activity for the period of the performance was undirected by the experimenter and was secured by means of the following written instructions: "At the conclusion of each test you are asked to dictate an introspection of your awareness of speed, or movement, or push, or drive, or motive. Start with the last awareness as you finish the test and work backward toward the beginning, recalling these awarenesses." Occasional promptings, "Is that all you recall?", were employed.

Conscious motivation was maintained at a high level, for a series of approximately 10 tests, throughout each of several days' observations. This was accomplished by administering systematically various new incentives, including knowledge of own record from test to test and knowledge of slightly superior records.

Observers. Four *Os* (*F, G, H, J*), worked in the major part of the experiment with tests of 1-min. duration.[1] One *O* (*I*) worked with tests of 4-min. duration.[2] All the *Os* were trained in reporting conscious activity. No specific training in the task preceded the work. The task consisted of Fryer's *Speed Addition Test*,[3] in which the problems are a chance distribution of totals of less than nine, arranged in rows of thirty.

Rate of performance is used as the criterion of validity. As indicated elsewhere,[4] performance has been found to follow conscious stimulation wherever intention can be specifically determined—in fact, far more closely than environmental stimulation and perhaps exactly within the range of capacity of the performer.

The descriptions of conscious stimulation averaged approximately 100 words, with very few totaling less than 50 words, which appeared to be as full undirected descriptions as could be gathered. There were 237 introspective descriptions distributed as follow: *F*, 46; *G*, 41; *H*, 32; *J*, 62; and *I*, 56 (4-min. tests)

Rating scales. The introspective descriptions were rated by *E* on two scales, Scale

* Accepted for publication September 22, 1938.

[1] Acknowledgement is made to Professor F. C. Bartlett, Drs. P. E. Vernon and M. G. Vernon and Mr. G. C. Grindley of the Psychological Laboratory of the University of Cambridge, England, who worked as the observers in the one-minute tests.

[2] The author, who wrote his own introspective reports. This work was performed prior to the main part of the experiment with the 4 *Os*.

[3] Douglas Fryer, *op. cit.*, 1936. Published by the Department of Psychology, New York University.

[4] Fryer, A genetic study of motivation under changing auditory situations, *Brit. J. Psychol.*, 24, 1934, 408-433; Motivation effects of auditory timing upon repetitive mental work, *ibid.*, 25, 1934, 140-169; Specific conscious intent and its correlates in performance, *ibid.*, 27, 1937, 364-393; Conscious activity in coördination of repetitive mental work and rhythmic timing, *ibid.*, 28, 1937, 150-166; Variability in automatic performance with uniform intent, *J. Appl. Psychol.*, 21, 1937, 528-545.

A providing for a general estimate of conscious motivation and Scale B including several specific estimates of factors thought to exist in conscious stimulation. These scales are given below as used in the experiment.

In application of these scales a judgment upon the part of the rater was involved between what is irrelevant in the introspective report of awareness and what is a part of the working situation.

A. Rating Scale of Efficiency of Intention in Realization of Goal

To decrease performance				To increase performance		
tremendous	definite	slight	neutral	slight	definite	tremendous
3—	2—	1—	0	+1	+2	+3
1	2	3	4	5	6	7

A generalized rating of efficiency in operation of the various factors of conscious motivation (including focus of cues, clearness of goal, amount of intentional effort, etc.) in realization of goal.

B. Rating Scales of Conscious Factors in Intention

I. Scale for conscious effort or degree of intentionally exerted motor pressure (tension). Rating of the amount of intended effort or energy, as reported in verbalizations or kinesthetic and organic sensitivities, or in generalized statements of motor attitude or set. Restraining effort (motor inhibition) is rated as minus.

no information	none	slight	definite	tremendous
X	0	1	2	3

II. Scale for degree of focus of intentionally used cues. Rating of the clearness of an intentionally used cue, or cues, as reported in imagery, verbalizations, or in generalized statements. Rating includes any quantification of the cue, or cues, consistently used.

no information	none	vague	intermittently or fairly clear	consistently or completely clear
X	0	1	2	3

IIa. Scale for degree of focus of exerted motor pressure (tension). This additional rating is of the clearness with which motor pressure as a cue is a part of conscious intention.

no information	none	vague	intermittently clear	consistently clear
X	0	1	2	3

III. Scale for frequency (or number) of cues occurring in working situation. Rating of number of cues during a test, as reported in imagery or in verbalizations, or in generalized statements. Cues need not be used, but must be related to working situation.

no information	none	few (1 or 2)	several (3 or 4)	exceptional number (more than 4)
X	0	1	2	3

Results. Table I shows all correlations (rho) by *O* and observation between the criterion of validity, rate of performance, and the various ratings of conscious motivation in the rating scales. Performance was computed for this analysis as percentage change from previous test based on Test 1 of each observation as 100%. Table II gives the arithmetic means and Table III the absolute means of correlations. The absolute means indicate the extent of correlation in either direction and the arith-

metic means the extent of correlation in the direction accepted by the experimenter as positive conscious motivation.

Only fair validity is indicated for the method of post quantification of the undirected introspective description, as used here. The average (arith.) correlation of the one-minute tests for Scale A, the general scale, is about 0.30. This correlation is raised only a few hundredths to 0.36 with the best assembling of specific ratings of Scale B, with 0.44 as the highest average (arith.) coefficient for any one rating.

TABLE I

CORRELATION (RHO) BETWEEN QUANTIFIED INTROSPECTIONS AND PERCENT CHANGE FROM PREVIOUS PERFORMANCE BY DAILY OBSERVATIONS

(Daily observations of approximately 10 tests.)

O	Observations	Quantified introspections						
		A	B_I	B_{II}	B_{IIa}	B_{III}	$B_{I+II+IIa}$	$B_{I+II+IIa+II}$
	I	0.30	0.48	0.57	0.14	0.10	0.43	0.36
F	II	0.10	−0.11	0.40	−0.06	0.28	−0.02	0.04
	III	0.44	0.05	0.78	0.05	−0.37	0.36	0.27
	IV	0.25	0.03	0.45	0.78	−0.17	0.50	0.38
	I	0.61	0.32	0.50	0.39	−0.40	0.38	0.34
G	II	−0.15	−0.02	0.75	0.41	0.00	0.52	0.50
	III	0.25	0.65	0.78	0.58	−0.37	0.78	0.85
	IV	0.48	0.43	0.61	0.42	0.26	0.57	0.89
	I	0.23	0.21	−0.26	−0.14	0.10	−0.12	−0.22
H	II	0.48	0.45	0.67	0.33	−0.08	0.46	0.45
	III	0.53	0.63	0.32	0.30	0.05	0.44	0.37
	I	0.02	0.35	0.10	0.01	−0.36	0.13	0.04
	II	−0.18	−0.02	0.01	0.28	0.06	−0.15	−0.20
J	III	0.06	0.34	0.05	0.29	0.08	0.26	0.16
	IV	0.42	0.27	0.44	0.39	−0.24	0.46	0.30
	V	0.81	0.43	0.83	0.68	0.50	0.79	0.80
	VI	0.34	0.35	0.46	0.10	0.50	0.32	0.32
	I	0.47	−0.07	0.47	0.50	−0.29	0.31	0.12
	II	0.49	0.35	0.11	−0.11	0.33	−0.12	0.03
	III	0.11	0.17	0.08	0.34	0.41	0.04	0.24
I	IV	0.13	−0.55	0.10	0.68	−0.04	−0.13	−0.17
	V	0.27	0.16	0.61	0.37	0.31	0.69	0.73
	VI	0.06	−0.16	0.12	0.19	0.20	−0.06	−0.01
	VII	0.26	0.22	−0.38	0.55	−0.14	0.24	0.01

It seems doubtful that further statistical treatment of these data would offer additional information as the conditions of validity in such an experiment usually lie within the data where no statistical treatment can correct them.

Certain improvements in method are suggested in the experiment which may raise the validity of post quantification of introspective data to a degree satisfactory for exploratory purposes in the study of conscious motivation. It is suggested that the descriptions of conscious activity for one-minute tests are more valid than for four-minute tests. In recent unpublished research by the author it has been found possible to secure reliable measures of 10 second performance with the same test material as used in this experiment, but the observers must be trained in accurate starting and stopping of work. A further trial of method should be worth-while with 10 second tests where more detailed descriptions of conscious activity are possible.

The author is interested in this method because of its possible application in the technical occupations. Intelligent workers may be able to report conscious activity for correlation with parallel objective measures, such as motion and time-study records. In such applied situations intensive training in introspective technique is

TABLE II

MEAN (ARITH.) OF CORRELATIONS (RHO) FOR DAILY OBSERVATIONS BETWEEN QUANTIFIED INTROSPECTIONS AND PERCENTAGE CHANGE FROM PREVIOUS PERFORMANCE

O	No.	Quantified introspections						
		A	B_I	B_{II}	B_{IIa}	B_{III}	$B_{I+II+IIa}$	$B_{I+II+IIa+III}$
F	4	0.27	0.11	0.55	0.23	−0.04	0.33	0.26
G	4	0.30	0.35	0.66	0.45	−0.13	0.56	0.65
H	3	0.41	0.43	0.24	0.16	0.02	0.26	0.20
J	6	0.25	0.29	0.32	0.29	0.09	0.30	0.24
Mean		0.31	0.30	0.44	0.28	−0.01	0.36	0.34
I	7	0.26	0.02	0.16	0.36	0.11	0.14	0.14

impractical and the undirected report offers a possible source of as yet unaccessible psychological knowledge of a variety of skills.

Suggestions. Several interesting suggestions in the study of conscious motivation are indicated in this analysis of method while the data are insufficient for a satisfactory statistical treatment of them.

(1) The B_{II} rating, which is the "Scale for degree of focus of intentionally used cues," has the highest correlation with the 1-min. performance and may carry

TABLE III

MEAN (ABSOLUTE) OF CORRELATIONS (RHO) FOR DAILY OBSERVATIONS BETWEEN QUANTIFIED INTROSPECTIONS AND PERCENTAGE CHANGE FROM PREVIOUS PERFORMANCE

O	No.	Quantified introspections						
		A	B_I	B_{II}	B_{IIa}	B_{III}	$B_{I+II+IIa}$	$B_{I+II+IIa+III}$
F	4	0.27	0.17	0.55	0.26	0.23	0.33	0.26
G	4	0.37	0.36	0.66	0.45	0.26	0.56	0.65
H	3	0.41	0.43	0.42	0.26	0.08	0.34	0.35
J	6	0.31	0.29	0.32	0.29	0.29	0.35	0.30
Mean		0.34	0.31	0.49	0.32	0.22	0.40	0.39
I	7	0.26	0.24	0.27	0.39	0.25	0.23	0.19

the full load of the ratings. This is interesting as it points to conscious focus of intentionally used cues as an important factor in increasing performance. This suggestion is supported in a previous study where observers were working with intention to maintain rate without change.[5] The correlation between estimated degrees of focus of intentionally used cues and variability in performance in reverse order was 0.84 ± 0.08.

(2) The B_{IIa} rating, which is the "Scale for degree of focus of exerted motor pressure (tension)," has a higher correlation with 4-min. performance than with 1-min. performance, which suggests that intentional motor pressure becomes more essential in highly motivated activity during lengthy periods of performance. The

[5] Fryer, *op. cit., Brit. J. Psychol.*, 27, 1937, 364-393.

literature as to the importance of motor effort in performance is contradictory and it may be that the time factor has been neglected.

(3) The $B_{II\ 1}$ rating, which is the "Scale for frequency (or number) of cues occurring in the working situation," offers the lowest correlation with the criterion. Seven of 17 coefficients for the 1-min. tests and 3 of 7 for the 4-min. tests are negative. Here is a suggestion that a large number of conscious cues is a disturbing factor in immediate performance.

THE EFFECT OF THE IMMEDIATE TEST ON VERBATIM AND SUMMARY RETENTION

By ALLEN L. EDWARDS and HORACE B. ENGLISH, Ohio State University

Striking systematic differences in retention have been found, in a series of studies of substance learning,[1] which seem to depend in the first instance upon the method of measurement employed. With the passage of time, recognition of sentences which are drawn verbatim from the text studied shows a significant loss similar to the Ebbinghaus curves of forgetting, while recognition of statements which give the substance of paragraphs shows little or no loss—in many cases a net gain or reminiscence effect.

The divergence between the curves of retention is a fact, and can be produced at will, though by no means without considerable effort in devising the measuring instruments. The persistence of the divergence under variation of a considerable number of factors has been investigated with approximately 3,000 *Ss* and in eleven separate experiments. This report deals with the effect of an immediate test upon the retention as measured in the two ways previously referred to. That is, does the rehearsal afforded by a test immediately after learning influence the two types of items in the same way and to the same extent, or does it operate differently upon the two?

Experimental procedure. A total of 155 college students were divided into 6 groups, 5 experimental and 1 control, each group having from 20-30 students in it. These groups were fairly homogeneous in respect to the learning abilities involved, being randomly selected from a single class.[2] Until much more is known about these abilities and their correlations, any rigorous attempt at matched groups is more likely to introduce, than to eliminate, systematic errors.[3] Each of the five experimental groups read and heard read an unfamiliar passage. Group A, given the immediate test after the reading, was retested at 60 and again at 90 days. None of the other experimental groups was given the immediate test. Instead one group, Group B, was tested at 5 days after the reading, another, Group C, at 10 days, Group D at 15 days, and Group E at 20 days. The Control Group which had not heard the passage read at all was given the test twice with an interval of 30 days separating the first and second tests.

The test, with slight modifications, was the same as that used by English, Welborn, and Killian in some of the earlier reports of *verbatim* and *summary* learning at the college level[4] and is therefore not reproduced here. The split-half reliability (corrected) is $0.82 \pm .02$.

---

* Accepted for publication August 12, 1938. This study was presented at the Madison, Wisconsin, meeting of the Midwestern Psychological Association on April 23, 1938. This is the twelfth of a series of studies in substance learning and retention.

[1] H. B. English, E. L. Welborn, and C. D. Killian, Studies in substance memorization, *J. Gen. Psychol.*, 11, 1934, 233-260; H. B. English, Further data on reminiscence, *Psychol. Bull.*, 32, 1935, 688.

[2] Acknowledgement is made to Professor J. E. Janney of Western College who administered some of the tests of his classes.

[3] Cf., on this point, English, The procedure of matched cases—a caution, *J. Educ. Psychol.*, 30, 1939, (in press).

[4] English, Welborn, and Killian, *op. cit.*, 238.

Results. Fig. 1 shows the trend of retention in the two kinds of items. The V's stand for mean verbatim scores; the S's for corresponding mean summary scores. The mean scores of the A Group for immediate, 60-, and 90-day tests are connected by a *straight line;* the mean scores for the two tests of the Control Group are connected by a *dotted line;* and the mean scores of the 5-, 10-, 15-, and 20-day tests (Groups B, C, D, and E) are connected by a *broken line.*

On the assumption that our groups are equivalent—which is subject to the previously mentioned qualifications—the points connected by the broken line represent a retention curve with *elimination of the rehearsal element* involved in successive testings. It can be easily seen that there is a sharp drop in the V-scores and

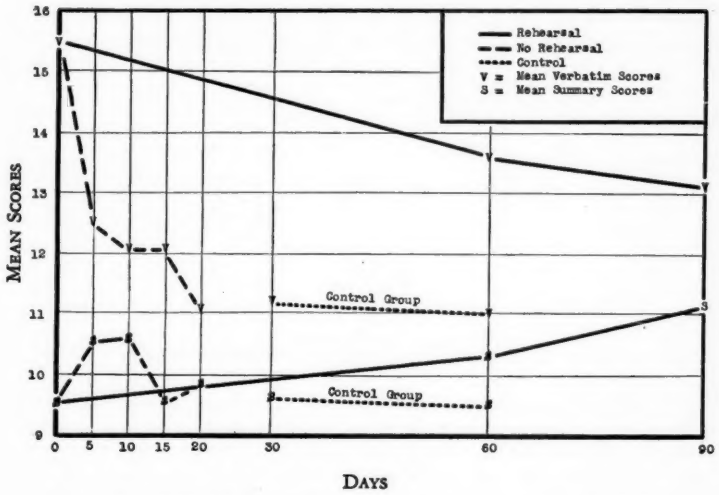


FIG. 1. RETENTION CURVES OF V- AND S-ITEMS WITH AND WITHOUT THE IMMEDIATE TEST

that this is not paralleled at all in the S-scores. For the V-scores, the decline from A (the immediate test) to B (the 5-day test) is statistically significant; the other successive declines, B to C, C to D, and D to E, do not meet the test of significance, perhaps because the populations are small and the P.E. correspondingly large. The total trend is, however, unmistakable and is wholly in line with other experimental data in this series and in other studies.

Turn now to the same comparison of the S-scores. In all cases there is found a gain in the more remote tests over the immediate. While these gains do not meet the requirements of significance, the important point is the difference between the trend in the V-scores and the S-scores—one definitely downward and the other upward, or at least not down. This difference is obviously significant statistically.

Neither the mean S-score nor the mean V-score for the Control Group showed any significant fluctuation or practice effect from the first to the second test. There

is thus one answer: the divergent trends found in the previous experiments are not due wholly to the rehearsal involved in testing immediately after the learning period. There is better net retention—less forgetting, more reminiscence, call it what you will—in *summary* than in *verbatim* even without the immediate test.

None the less the rehearsal afforded by the immediate test is not without its effect. Compare the two *verbatim* curves, the one with and the other without the rehearsal. Both trends are unambiguously downward, but the trend is much sharper where the immediate test is lacking. Or take the two *summary* curves; the trend is slightly upward without the rehearsal, but with it it is somewhat more so.

Table I gives the essential statistics on this point. The difference between the means of the immediate and the 60-day test for the V-scores has a critical ratio of 4.1; for the 90-day loss it is 4.9. The S-scores, on the other hand, show a gain; instead of forgetting there is reminiscence. The gain from the immediate to the

TABLE I

COMPARISON OF THE MEANS OF V AND S ITEMS OVER 90 DAYS
(Group A: N = 30)

Mean Scores	Test periods			
	Immediate	60-day	90-day	Diff./P.E.$_{diff.}$
V	15.5±.25	13.8±.32		4.1
		13.8±.32	13.2±.40	1.2
	15.5±.25		13.2±.40	4.9
S	9.5±.33	10.3±.32		1.7
		10.3±.32	11.1±.32	1.8
	9.5±.33		11.1±.32	3.5

60-day test is not significant but for the entire 90-day period the gain of 1.6 has a critical ratio of 3.5.

These findings indicate rather conclusively that rapid forgetting takes place in the V-items, but is somewhat retarded by an immediate test. In the S sub-test, forgetting is largely replaced by reminiscence. It is difficult to say from the data just what the effect of the immediate test is on the S-items. The S-curve is still rising, though slowly, at the 20-day interval even without the immediate test. It is apparent, however, that the immediate test is not nearly so important a factor in determining the rise in the S-curve, as it is in slowing the downward drop of the V.

It is perhaps worthwhile to emphasize that apparently it is the *immediate* rehearsal which counts. English, Welborn, and Killian found that the general form of his retention curves was but little altered by the dropping out for some of their groups of test-reviews at 8-, 14-, and 30-days. Under comparable experimental conditions with the same test material, we find a considerable change when we drop out only one review—the immediate test.

Discussion. Serious consideration must thus be given the value of relatively immediate testing after study. It is impossible to compare the retention of the A Group, which was given the immediate test, and that of the B, C, D, and E Groups, which were tested at intervals from 5 to 20 days after study, without being impressed by the very great value of some sort of immediate review. Does it have no significance for the teaching situation in our schools and colleges that the immediate review which a test provides—probably better than any other device—is so important in

retarding forgetting? Yet, in how many cases is practical heed paid to this fact? It is a fair deduction that lectures and recitations should be closed in time to allow students a brief review or quiz. Students should be more strongly urged to intersperse reviews with their study.

The theoretical implications may be more briefly dealt with. The divergence in retention previously reported does not disappear with the alteration of still another variable. Whatever the interpretation of the divergence, therefore, we believe it is not an artifact of the experimental procedure. We have added one more straw at least to the burden of proof which must be borne by those who defend a qualitatively single learning process.

THE VISUAL PERCEPTION OF FILLED AND UNFILLED SPACE BY THE AFTER-IMAGE METHOD

By ALFRED L. BALDWIN, Harvard University

Koffka has suggested that the stability of the visual field is reduced in the after-image.[1] This view is substantiated by the works of Noll[2] and Rothschild.[3] The results suggest a technique for testing some of the implications of the theory of perception of movement presented by Brown and Voth,[4] and extended to the perception of stable configurations by Orbison.[5] The theory in its present form will not predict the results of the studies of the apparent size of filled and unfilled space. Knox[6] and more recently Spiegel[7] find that the phenomenal distance between two points is least when the interval between them is empty. As the interval is gradually filled by adding new dots to the figure, the phenomenal size increases to a maximum value. Further addition of dots decreases the apparent length of the interval, and the continuous line, while phenomenally longer than the unfilled interval, is phenomenally shorter than the maximum value.

The use of the after-image in the study of this problem seems promising for two reasons. (1) The after-image of an objective situation is distorted in the same direction as the phenomenal appearance of the objective situation. (2) The distortion of the after-image can be objectively measured.

As a first step in the use of this method the writer has undertaken a brief experimental investigation of the after-images of filled and unfilled space. The validity of the method may be tested by its agreement with the results of Spiegel, and the validity of the concept of 'cohesive force' for prediction of the distortion of after-images may be determined.[8]

EXPERIMENTAL PROCEDURE

A light-stimulus of a horizontal extent was presented to *O*. After observation for a sufficient length of time, a screen was lowered in front of the stimulus so that the

* Accepted for publication October 27, 1938.

[1] K. Koffka, *Principles of Gestalt Psychology*, 1935, 138, 141.

[2] A. Noll, Beiträge zur Psychologie der Gestalt: XIII. Versuche über Nachbilder, *Psychol. Forsch.*, 8, 1926, 3-28.

[3] H. Rothschild, Ueber den Einfluss der Gestalt auf das negative Nachbild ruhender visueller Figuren, [Graefes] *Arch. f. Ophthalm.*, 112, 1923, 1-28.

[4] J. F. Brown and A. C. Voth, The path of seen movement as a function of the vector-field, this JOURNAL, 49, 1937, 543-563.

[5] W. D. Orbison, Shape as a function of the vector-field, this JOURNAL, 52, 1939, 31-45; also 309.

[6] H. W. Knox, On the quantitative determination of an optical illusion, this JOURNAL, 6, 1894, 413-421.

[7] H. G. Spiegel, Ueber den Einfluss des Zwischenfeldes auf gesehene Abstände, *Psychol. Forsch.*, 21, 1937, 327-382.

[8] The experimental work for this study was done in the Psychological Laboratory of the University of Kansas. The writer wishes to thank Dr. J. F. Brown of the University of Kansas and Dr. E. G. Boring of Harvard University for their assistance in preparing this report.

projection-surface was practically the same distance from *O* as the stimulus. On the screen were two pointers which *O* was instructed to adjust until they were at the ends of the interval in the after-image. On the end of each pointer was a small point of light to make the pointers readily visible.

The measurement is made by superimposing a stimulus-field (two points of light used to measure the interval objectively) upon a field of after-image (two or more spots of light). The success of the method depends upon the fact that these two superimposed fields act, each as a whole and each relatively independently of the other. Thus, when the termini of the after-image, moving under field forces, expand or contract the interval, they do not carry the previously coincident perceived points of light with them. These perceived points, being anchored to their stimuli, can be moved to coincide with the changed after-image only by moving their stimuli. It is for this reason that the movement of the stimuli can be used to measure the change in the after-image that occurs under its field forces. This method has been described by Koffka and found to be satisfactory.[9]

In order to assist *O* in maintaining fixation, observations were made through artificial pupils.

Two *O*s (*A* and *B*) were used. They were undergraduates with no psychological training and were naïve with regard to the purpose of the experiments. They were given two weeks preliminary training in the observation of after-images of single points of light. *A* gave results that were more consistent than those of *B*. He also experienced less difficulty in making the observations.

Results

There were five experiments, each of which consisted of 50 observations.

Experiment 1. In this experiment *O* adjusted one point to the after-image of a single point of light 8 cm. to the right of the fixation-point. In a second situation the stimulus was 8 cm. to the left of the fixation-point. These measurements were designed to determine as nearly as possible the perceptual positions of the lights independently of all field forces. The results in the first part of Table I show the two halves of the experiment combined to give the theoretical length of an objective 16 cm. interval. The two *O*s differed rather widely under this condition, but the standard deviation for each is quite small.

Experiment 2. Two lights were presented simultaneously at the same positions as in Experiment 1. According to the Brown-Voth theory the distance between the after-image of two lights should be less than it was in Experiment 1 because of the cohesive force between the two lights. The results in the second part of Table I confirm this theory. For both *O*s the distance between the points of light was decreased, although the decrease was much more marked for *A* than for *B*.

Experiment 3. Five evenly spaced points of light covering the same distance were presented. The theory predicts that the distance will decrease still further unless there should be an increase in 'restraining force' which would cause the interval to approach its objective length. Spiegel found that the interval appears longer under these conditions than under the conditions of Experiment 2. The results in the third part of Table I conform to Spiegel's results. In the case of the more consistent *O*

[9] K. Koffka, Ueber die Messung der Grösse von Nachbildern, *Psychol. Forsch.*, 3, 1923, 33-41.

(*A*) this increase was enormous. An objective distance of 16 cm. was observed to be 37.5 cm! So large a difference needs checking, but a repetition of the situation for 50 more observations left the results essentially unchanged. The theory is unable to predict this increase.

Experiment 4. A line 16 cm. long was presented. Spiegel's results would indicate that the line should seem to be shorter than the five-dot situation described in Experiment 3. Again the theory would be unable to account for apparent lengths greater than those obtained in Experiment 1. The results of *A* shown in the fourth part of Table I confirm the results of Spiegel's experiment. The results of *B* in this

TABLE I

THE DISTANCES OF THE POINTERS FROM THE ENDS OF THE OBJECTIVE INTERVAL AND THEIR DISTANCES FROM EACH OTHER

A positive value indicates that the pointer is placed further from the center than the end of the objective interval; a minus value, that it is placed closer. The results under "Combined" were obtained by adding 16 cm. (the objective length) to the sum of the values of the two pointers. Every entry is the average of 50 adjustments.

Experiment No.	O	Left pointer		Right pointer		Combined	
		cm.	σ	cm.	σ	cm.	σ
1	A	2.4	.9	3.9	.8	22.3	1.2
	B	2.0	.8	−3.7	.7	14.3	1.0
2	A	−4	.6	1.0	.4	13	.8
	B	4.7	1.2	−7.4	1.0	13.3	1.6
3	A 1st 50	16.7	.8	9.7	.6	42.4	1.0
	A 2nd 50	9.2	.9	8.6	1.0	33.8	1.3
	A av.	13.4	3.9	9.1	.9	38.5	4.0
	B 1st 50	1.4	1.0	−1.6	1.0	15.8	1.5
	B 2nd 50	5.8	1.1	−4.9	.8	16.9	1.3
	B av.	3.6	2.3	−3.2	1.8	16.4	2.9
4	A 1st 50	5.0	.9	5.0	.9	26.0	1.3
	A 2nd 50	5.0	.9	11.0	.8	32.0	1.2
	A av.	5.0	.9	8.0	3.1	29.0	3.2
	B 1st 50	11.4	1.1	8.2	.9	35.5	1.4
	B 2nd 50	−1.0	.8	1.0	.4	16.1	.9
	B av.	5.2	6.6	4.6	4.1	25.8	7.9
5a. Repetition of 3	A	8.2	1.1	16.2	1.2	40.4	1.6
	B	−3.6	1.0	−3.8	.8	8.6	1.3
5b. Repetition of 4	A	9.0	.7	16.0	.7	41.0	.9
	B	−3.5	1.0	.1	.5	12.6	1.0

case do not correspond. In addition, the large difference between the two series of trials shows the results of *B* to be unreliable.

Experiment 5. This experiment is a repetition of Experiments 3 and 4 without the use of artificial pupils. *A* shows no marked change under the conditions of Experiment 3, but the repetition of Experiment 4 shows a significant increase. *B* shows a marked decrease in the perceived distances. These results are inexplicable at the present time.

DISCUSSION

The effect of the data is to confirm the results of Spiegel and to embarrass the theory of Brown and Voth as it is presented in their report. We should note, however, that the work of Brown and Voth and of Orbison is for conditions which

resemble the situation of Experiment 2. In nearly every case these investigations were concerned with the interaction of two objects. Even the experiments reported by Brown and Voth in which there were four objects in the visual field never exhibit more than two objects in a line and are thus reduced to the resultant of two situations in which there are only two objects acting at a time. In the same way Orbison's data are concerned with the interaction of one line up on another, and not with the apparent size of the whole field. It would seem likely, then, that the forces described for the prediction of such phenomena do not predict for other types of situations. What must eventually be done is describe these dynamic systems in more general terms so that the same forces will be 'cohesive' in certain situations and 'expansive' in others.

A notable result of the present experiment is to show that the changes in the after-images under different conditions may be enormous. The distortion that ordinarily occurs in the perception of filled and unfilled space is of the order of 5 to 10% of the length of the line, whereas with after-images one may obtain increases of 100 or even 200%.

SUMMARY

(1) Measurements were made of the after-images of filled and unfilled space.

(2) The data correspond to Spiegel's results in exhibiting first a shrinking and then an expansion as the interval is filled. They do not exhibit the final shrinking as the interval is made continuous.

(3) The data confirm the theory of Brown and Voth in predicting a shrinking, when a situation showing no cohesive forces is changed to one in which cohesive forces are present, *i.e.* from Experiments 1 to 2.

(4) The distortions recorded by this method may be much greater than those obtained in the usual study of illusions of filled and unfilled space.

THE SEX DIFFERENCE ON THE COLOR-NAMING TEST

By PHILIP H. DUBOIS, University of New Mexico

It has frequently been observed that women and girls reply in general more quickly than men and boys when tested with the Woodworth-Wells color-naming test. Ligon presented results based upon the records of 638 *Ss*, ranging from 6 to 18 yr. in age and from the first to the ninth grade in educational placement.[1] He also used the word-reading test in which the names of five colors were read from typewritten sheets. From the fact that the differences were greater and statistically more reliable with the color-naming test than with the word-reading test, he argues that the sex difference on the color naming test is based on a "color naming special factor," rather than on verbal facility.

Ligon was interested only incidentally in the sex difference, as his experiment, like most of those reported in which the color-naming test is used, was chiefly concerned with the effect of practice.

The present experiment represents an investigation of Ligon's hypothesis. The *Ss* were 149 men and 133 women, all junior college students, mostly of English and Scandinavian extraction. Records of several other *Ss* who did not show normal color vision on the Ishihara test were discarded.

Three tests were used: the Woodworth-Wells color-naming test with oral responses, the word-reading test, and a non-verbal form of the Woodworth-Wells test. For the latter a board was prepared on which was pasted a rectangle, 3 × 5 in. in size, of each of the following five colors: red, green, yellow, blue, and black. *S* indicated his response by tapping with a stylus on the rectangle corresponding to the stimulus-color. Every test, involving 100 responses, was given twice at a single sitting, but in the two trials the order of responses was different. The sequence of tests was: color-naming A; word-reading A; non-verbal A; non-verbal B, word-reading B, and color-naming B. Scores were recorded in terms of the number of seconds necessary to complete each test. When an error was made *S* was required to correct it. The number of errors was negligible.

TABLE I

CORRELATIONS OF THE TWO TRIALS OF EVERY TEST

Test	Men (N=149)		Women (N=133)	
	Obt.	Sp. Br.	Obt.	Sp. Br.
Color-naming	.56	.72	.83	.91
Word-reading	.80	.89	.86	.93
Non-verbal	.88	.93	.88	.93

Reliabilities of the tests were computed by correlating scores on the first and second trials of each test. Since intercorrelations were computed using the sums of the scores on the two trials, reliabilities are also presented as estimated by the Spearman-Brown formula. Reliabilities corrected to two decimals are given in Table I.

* Accepted for publication December 10, 1938.

[1] E. M. Ligon, A genetic study of color naming and word reading, this JOURNAL, 44, 1932, 103-122.

The higher reliability of the non-verbal test may have resulted from the fact that the two trials followed one after the other, while the color-naming test appeared at the beginning and end of the series.

Table II gives means, standard deviations, differences between means of men and women and critical ratios ($D/S.D._{diff.}$) for the two trials and for summed scores on each test.

It will be noted that differences on all tests favor the women in that they have lower scores. On the first trial the difference on the color-naming test is to be considered significant according to the conventional standard. On the word-reading test the differences cannot be considered completely reliable. The greatest differences and the differences most reliable statistically are those on the non-verbal form of the Woodworth-Wells color-naming test. This test involves a double response to color; first on the stimulus-sheet, and second in making the choice among the rectangles of color. A plausible interpretation is that the non-verbal test involves Ligon's "color-naming special factor" to a double degree. Perhaps this factor should be re-christened a "color recognizing special factor." The results give clear-cut support to the hypothesis that the sex difference on the color-naming test is not to be attributed to differences in verbal facility but are due to the operation of a special factor.

TABLE II

THE PERFORMANCES OF THE MEN AND WOMEN IN THE DIFFERENT TRIALS WITH EVERY TEST

Test	Men (N=149)		Women (N=133)		D	$D/S.D._{diff.}$
	Mean	S.D.	Mean	S.D.		
Color-naming A	66.60	11.19	62.25	10.80	4.35	3.31
Color-naming B	63.51	9.72	62.04	11.27	1.47	1.17
Total	130.12	18.46	122.18	20.99	7.94	3.36
Word-reading A	43.48	7.09	41.49	7.55	1.99	2.27
Word-reading B	39.96	7.18	38.40	6.17	1.57	1.97
Total	83.44	13.53	79.86	13.08	3.59	2.26
Non-verbal A	89.90	11.95	80.87	12.49	9.03	6.19
Non-verbal B	87.41	12.75	79.71	13.04	7.70	5.00
Total	177.32	23.94	160.14	24.86	17.18	5.90

The evidence presented in Table II is confirmed from a study of the medians. Medians of total scores are given in Table III.

TABLE III

MEDIANS OF THE TOTAL SCORES OF THE VARIOUS TESTS

Test	Men	Women
Color-naming	127.65	119.35
Word-reading	81.86	78.53
Non-verbal	176.64	158.50

On the color-naming test, 36.7% of the men attained the median score of the women, compared with 40.6% on the word-reading test, and only 19.9% on the non-verbal test.

Intercorrelations. Intercorrelations of the three tests, and partial correlations with the variability due to the third test held constant are presented in Table IV.

There is some indication that the abilities measured by the tests are somewhat differently organized in men and women. The greatest discrepancy in zero order correlations is between word-reading and the non-verbal test. The difference in the two correlations is 0.22. The critical ratio of the difference ($D/S.D._{diff.}$) is 2.2,

TABLE IV

INTERCORRELATIONS AMONG THE THREE TESTS

Tests	Men		Women	
	Zero order r	Partial r	Zero order r	Partial r
Color-naming and word-reading	.51± .04	.36	.61± .04	.57
Color-naming and non-verbal	.53± .04	.34	.53± .04	.47
Word-reading and non-verbal	.49± .04	.31	.27± .05	−.07

showing that the chances are approximately 99 in 100 that there is a difference greater than zero.

With the men, partial correlation reduces the correlations somewhat, but all tests reveal some degree of relation. With the women, however, the correlation between word-reading and the non-verbal Woodworth-Wells test becomes slightly negative when the variability due to the color-naming test is held constant.

Brown and Thompson state that when for three variables $r^2_{ab} + r^2_{ac} + r^2_{bc} + 2r_{ab}r_{ac}r_{bc}$ is significantly greater than unity there must be a common factor running through the tests.[2] For the men, this value is 1.05, and for women, 0.90. There is some evidence, therefore, that the three tests show a common factor—possibly a speed factor—for men but not for women.

Conclusions. The results of this study lead to the following conclusions.

(1) Ligon's hypothesis of a "color-naming special factor" rather than a verbal factor accounting for the sex difference on the Woodworth-Wells color-naming test is confirmed.

(2) The sex difference is increased if non-verbal responses involving tapping on the appropriate color are substituted for the usual oral response. On this non-verbal test the difference between the sexes is striking, with only 19.9% of the men attaining the women's median.

(3) There is some evidence of difference in the organization of the abilities measured by the three tests in the two groups of *Ss*.

[2] W. Brown, and G. H. Thompson, *Essentials of Mental Measurement,* 1925, 142.

THE INTER-SOCIETY COLOR COUNCIL SYMPOSIUM ON COLOR TOLERANCE

INTRODUCTION

In the first Articles of Organization and Procedure of the Inter-Society Color Council there occurs the following statement: "The aims and purposes of the Council are to stimulate and coördinate the work being done by various societies and associations leading to the standardization, descriptions and specification of color, and to promote the practical application of these results to the color problems arising in science, art and industry." In furtherance of these aims, successive administrations of the council have sponsored a wide variety of activities. Problems of a concrete nature have been assigned to individuals and to committees, and attempts have been made to crystallize questions that were in a more nebulous state. In addition, it has been evident from the beginning that a mutual exchange of ideas performs a valuable function where questions are not sufficiently formulated. As a result, in the business meetings, committee reports have become veritable symposia. Miss Nickerson has shown, in her historical review of the activities of the Council, how more extensive programs were added to its annual meetings. There was at first, I believe, no clearly outlined policy. The executive Committee made use of available program material in an exploration of the desirability of expanding Council activities in that direction. The answer was unmistakable and last year Mr. Little demonstrated what could be done with a technical program built about a special topic. We have followed much the same plan this year, without the implication, however that we are establishing a permanent policy. It is to be expected that the organization will remain responsive to new and changing demands from workers in the field of color, and that its activities will evolve accordingly.

The present group of papers centers about a topic with which every person who deals with colors must eventually be concerned. That does not mean that the problem will appear in the same light from all angles. Indeed, the certainty that it would not made it appear especially suitable for sponsorship by the Council. The particular papers were invited with the expressed understanding that their function is not to present new technical data to workers in their special field, but to acquaint those in other fields with established facts and techniques. It was recognized that color tolerance is in large measure a psychological matter, and therefore the major call for contributions was made upon psychologists. Nevertheless, other aspects of the problem were not lost to sight. We may have been too ambitious, but it is our hope that by considering it from various angles, color tolerance may be seen in a new light from every point of view, and in this way old difficulties be resolved and new problems be formulated that will carry all of us a step forward.

Hobart College FORREST LEE DIMMICK, Chairman

* Accepted for publication February 1, 1939. These papers were presented at the eighth annual meeting of the Inter-Society Color Council in coöperation with the American Psychological Association, New York, February 23, 1939.

THE PSYCHOPHYSICS OF COLOR TOLERANCE

By EDWIN G. BORING, Harvard University

Any particular color tolerance can and must be measured in terms of the physical units of the color-stimulus, since the stimulus-scale is an ultimate system of reference. One hopes, nevertheless, for some general statement of the limits of any particular kind of tolerance, and it would seem that such a generalization must be more likely to be expressible in the units of some psychological scale which is definitely but perhaps not linearly related to the physical scale of the stimulus.

Tolerance is a matter of perception and must therefore be related to discrimination. There must be as many kinds of tolerance as there are reasons for fixing the character of colors, and two colors that are not noticeably different would necessarily lie within the range of all tolerances, if they are really not discriminable. That sentence is, however, much too simple to express the whole truth. While many tolerances are supraliminal, certainly there are others which are subliminal. The difference limen is the average jnd (just noticeable difference), that is to say, it is the difference which is as often perceived as not. A difference that is perceived only forty per cent of the time is subliminal, yet it is perceived almost, though not quite, as often as not. Complete intolerance, a requirement that two colors should never be perceived as different, would specify a difference that is far below the limen.

There are three psychological scales, any one of which might be used for measuring tolerance. (1) The first is the DL (difference limen). It can be chosen as the unit, and a tolerance specified as a permissible number of DL. (2) A more stable measure is σ (standard deviation) of the psychometric function. (3) The third measure is the sense-distance. Since these three measures are, unfortunately, almost the private property of the psychologists, I must undertake to explain them now, to show you their natures and disadvantages. My thesis is that, for the measurement of tolerances, the DL has the most disadvantages, and the sense-distance the least. Unfortunately it is the DL that has been used most often by psychologists, whereas the sense-distance has hardly been used at all.

(1) The difference limen. The physicist, when he encounters a problem of psychological measurement, generally works with the jnd. He changes one of two stimuli until the two impressions are just noticeably different, or else he decreases the difference until the two are just not noticeably different. Then he takes an average and calls it the jnd. I fear that sometimes he does not take more than four, or perhaps ten, readings, thus failing in proper respect for the great variability of the organism. Nevertheless in principle his procedure is correct: the average jnd is the DL.

Since this statement may not be immediately obvious to everyone, we had better validate it. Fig. 1 is a schematic diagram for showing why the average jnd is the point where a perception of difference is just as likely to occur as not.[1] The seven-step distribution curve is supposed to show how the organism varies in its sensi-

[1] E. G. Boring, A chart of the psychometric function, this JOURNAL, 28, 1917, 465-470; J. P. Guilford, *Psychometric Methods,* 1936, 172.

tivity to some specified kind of difference. When the organism is momentarily favorably disposed toward perceiving the difference, then a small difference will elicit a jnd. The heights of the steps of this curve are 1, 6, 15, 20, 15, 6 and 1—a total of 64 cases. That is to say, once in 64 times the organism will be so well disposed toward the perception of a difference that a small difference, represented by the stimulus-difference 8 will be perceived. If the stimulus-difference is as

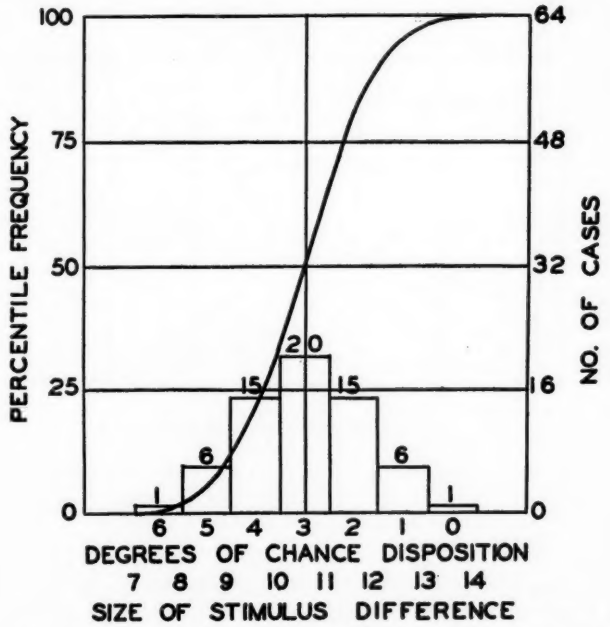


FIG. 1. RELATION OF THE PSYCHOMETRIC FUNCTION TO THE VARIABLE DISPOSITION OF THE ORGANISM

The degrees of chance disposition of the organism are shown in the seven-step distribution curve. The heights of these steps are 1, 6, 15, 20, 15, 6, 1—64 cases in all. The tall ogive is the psychometric function; it shows the sum from left to right of the levels in the other curve. The abscissa-scale shows the stimulus-differences that are candidates for being perceived. See text as to why the DL, the average jnd, lies at the midpoint 10.5. After E. G. Boring, this JOURNAL, 28, 1917, 466.

great as 9, then it will be perceived 7 times out of 64—the once when the disposition of the organism is what is called '6' in the diagram, and also the 6 times when the disposition is '5.' The tall ogive is the integral curve, the sum of the values on the distribution curve. It shows that at a middle value of the stimulus-difference—actually a value of 10.5—the difference will be sensed just as often as not, or the frequency of the perception of the difference will be 50%. This value is the DL and is the same as the average jnd. The ogive curve is called a *psycho-*

metric function, and the first advantage of its use is that it directs attention to the variabiliy that underlies all discrimination. When the physicist determines an average jnd he is likely to think of variability as an error to be got rid of by averaging, whereas it is, as we shall see presently, just exactly this spread of variability that measures discrimination.

Fig. 2 shows a typical set of psychometric functions for the perception of lifted weights.[2] The psychometric function for the judgement *greater* is a frequency

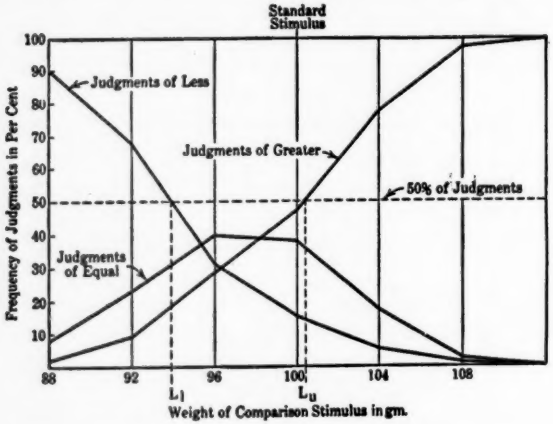


FIG. 2. PSYCHOMETRIC FUNCTIONS FOR LIFTED WEIGHTS
(A representative set. After D. W. Chapman in E. G. Boring, H. S. Langfeld, H. P. Weld et al., *Psychology: A Factual Textbook,* 1935, 55.)

function which shows how often the comparison weights were judged *greater* than the standard of 100 gm. There is a comparable function for the judgment *less.* The two curves do not, however, add up to 100% because the judgment *equal* was also allowed. The frequency of the use of the *equal* category is shown in the third curve. The lower and upper limens, L_l and L_u, are placed by definition at the 50% points. The upper limen is where the judgment *greater* is just as likely to occur as not; when it does not occur, the judgment may be either *less* or *equal.* Because of the use of the *equal* category, the two limens are not coincident, and their size is really dependent upon the frequency with which equality appears in the data. We shall return to this point in a moment.[3]

Now the real trouble with the DL as a measure of tolerance lies in the fact that its size depends upon the number of *equal* judgments, and that the use of that

---

[2] D. W. Chapman in E. G. Boring, H. S. Langfeld, and H. P. Weld, *Psychology: A Factual Textbook,* 1935, 55.

[3] The fact that, due to the 'time-error' of successive comparison, these particular curves are assymetrical about the value of the standard weight, 100 gm., need not concern us here, since most visual comparison is simultaneous.

category can not always be controlled. In Fig. 3 you see a hypothetical case where the range of equality is very large. The middle curve is for the *equal* judgments, and you see how its rise has forced the liminal points, L_l and L_u, apart. On the other hand, if the *equal* judgments fall off as they do in the low curve in Fig. 4,

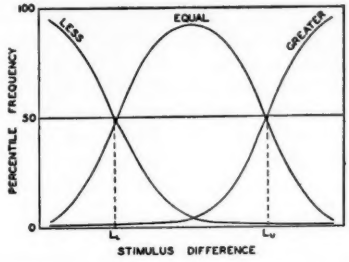


FIG. 3. PSYCHOMETRIC FUNCTIONS WITH LARGE EQUAL CATEGORY AND LARGE LIMENS

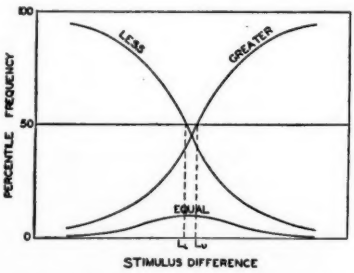


FIG. 4. PSYCHOMETRIC FUNCTION WITH SMALL EQUAL CATEGORY AND SMALL LIMENS

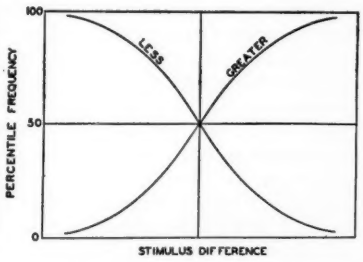


FIG. 5. PSYCHOMETRIC FUNCTIONS WITH NO EQUAL CATEGORY AND VANISHED LIMENS

the limens get very small. If the judgment of equality is eliminated entirely, as in Fig. 5, then the limens also disappear entirely. The two remaining psychometric functions, one for *greater* and one for *less*, now add up to 100%: either entirely determines the other.[4]

It has been found that the number of *equal* judgments can be altered by changing the attitude of the observer toward this middle category of judgment. In some cases the observer may use the category quite liberally, but then it becomes possible to diminish its use or even to abolish it altogether. If you say to an observer: "Of course, you know that no two stimuli can ever be exactly the same; there is always some difference; try your best to see the difference,"—if you say that, then the observer will give less *equals* and more *greaters* and *lesses* and the psychometric

[4] Cf. in general Guilford, *op. cit.*, 194, 196.

functions may look like Fig. 4. If you say: "Since the stimuli are never really equal, you must guess the difference; the fact is that people more often guess right than wrong under such circumstances,"—if you say that, then you are likely to get the situation in Fig. 5, where the *equal* category has disappeared and the limens are zero.

Now I have not made up this psychophysical trouble out of my imagination. It is a well-established fact that the attitude toward the use of the *equal* category varies somewhat adventitiously and is not easily controlled. Fernberger showed how instructions will diminish or abolish the category. Its rejection has been advocated ever since 1884 and the standard handbooks now point out the trouble.[5]

(2) *The standard deviation.* Even when the limens disappear as in Fig. 5, there is left a measure of sensitivity. It is represented by the spread of these psychometric functions. We can use as a unit the semi-interquartile range (half the distance from 25% to 75%), the probable error (which is the same thing in a symmetrical case), the standard deviation (σ), or the modulus for a normal distribution (h).[6] These measures are all related, and we shall speak of σ as the unit. The great advantage of the use of σ is that it is free of the uncontrolled *equal* category. We tell the observer that there is a difference whether he is sure of it or not, that he must observe it or guess it, and then we measure the spread of his responses by σ. In it we have a more constant and reliable unit for sensory discrimination than we had in the DL.

It has been argued that σ and the DL vary together, and that there is therefore no reason why σ should not be substituted for the DL. Such a relationship as a generality is, however, manifestly impossible, since change of attitude toward the equal category may not affect σ and must affect the DL. Holway has recently shown that for certain special conditions the DL for visual brightness varies with its σ when the level of intensity is changed, but this is a special case.[7] Although Holway did not explicitly control the attitude of his observers toward equality, he arranged his procedures so that it would be impossible for *equal* to disappear entirely. A different procedure might give very different values for the DL. At any rate the substitution of σ is desirable because it avoids the constant suspicion under which the DL labors.

It has been suggested that the use of σ may not be practicable in the case of simultaneous visual comparison of two apposed fields, for the reason that it might prove impossible to eliminate the *equal* category of judgment. Ordinarily the judgment *equal* in such a case involves the disappearance of the contour that separates the fields. Would the disappearance of the contour compel the judgment *equal?* We need experimentation to settle such a point, but it seems improbable that we

---

[5] For the history and references, see S. W. Fernberger, The use of equality judgments in psychophysical procedures, *Psychol. Rev.*, 37, 1930, 107-112; also E. Culler, Studies in psychometric theory, *Psychol. Monog.*, 35, 1926, 97-122; Guilford, *op. cit.*, 192 f., 201-205; R. S. Woodworth, *Experimental Psychology*, 1938, 421-425. I confess here to having once been on the other side of this argument: Boring, The control of attitude in psychophysical experiments, *Psychol. Rev.*, 27, 1920, 440-452.

[6] Culler, Guilford, and Woodworth, *locc. citt.*

[7] A. H. Holway, On the precision of photometric observations, *J. Opt. Soc. Amer.*, 27, 1937, 120-123.

have here to do with a special case. What is called a 'homogeneous' visual field never appears entirely homogeneous to the most careful inspection. It varies from moment to moment and from point to point, as the eyes move and adaptation takes place; and total field which has lost a dividing contour would still present a problem of decision between its two halves. There is little reason to expect that the *equal* judgment could prove compulsory here.

(3) The sense-distance. It is quite possible for an observer directly to compare two supraliminal sense-distances, and it would seem that such a measure must provide a correct unit for color tolerance. The DL and σ are measures of frequency of perceived difference. In questions of tolerance one is more likely to want to know how great the perceived difference actually is than how often the difference will be perceived. Wide limits of tolerance when measured in terms of the DL or σ are merely multiples of some small difference perceived with a given frequency, whereas measured in sense-distance a statement of tolerance has direct significance.

A sense-distance is determined by its two limiting points; hence it follows that a pair of sense-distances would be fixed by four points. Suppose that A, B, C and D represent four colors. Then an observer can be asked to say whether the color-distance from A to B is greater or less than the distance from C to D. The fact that this distance is empty does not necessarily make the judgment difficult. If A, B, C and D are four points marked in visual space, it is not hard to say whether the distance AB is greater or less than the distance CD, even though the lines AB and CD are not drawn in to fill the visual vacuum.

Actually there are available at least four methods for the measurement of sense-distances.

(a) The *method of equation* is the one I have just described.[8] Two pairs of limiting values are given and at least one limit is varied. Is AB greater or less than CD?

(b) The *method of bisection* is more usual.[9] In it three successive values are given, A, B and C. One of them, usually the middle one, B, is varied until the distance AB equals BC. This method is, of course, the method of equation with the two distances coterminous in the middle.

(c) In the *method of halving* there are only two stimuli. A variable stimulus is adjusted so that its sensation B appears to be half as great as a constant sensation A. Logically there is implicit in this procedure a third point, zero: B is made to bisect the distance between zero and A. This method is new but has been used successfully for the measurement of the loudness of tones[10] and also, surprising as that may seem, for the measurement of their pitch.[11] Other researches applying it to visual brightness and to times are said to be in progress.

(d) There is also possible a *method of fractionation*, which is like the method

---

[8] The method of equation is similar to the classical method of equivalents; cf. Guilford, *op. cit.*, 134 f.

[9] The method of bisection is essentially the classical method of equal sense-distances; see Guilford, *op. cit.*, 141-165; D. B. Judd, Saturation scale for yellow colors, *J. Opt. Soc. Amer.*, 23, 1933, 35-45, was in principle using this method.

[10] See S. S. Stevens and H. Davis, *Hearing*, 1938, 112-114, 117-120, and references there cited.

[11] See Stevens and Davis, *op. cit.*, 76-81, and references there cited.

of halving but uses fractions other than the half.[12] B may be adjusted so as to be one-third as great as A, or the stimuli may remain unchanged and the ratio of the distances between them estimated.

The first two methods are historically the oldest and they apply to different cases. The method of equation must be used in equating sense-distances that can not be coterminous, as, for example, the equation of a difference in lightness of two greens to a difference in brightness of two yellows. The method of bisection, being simpler, is desirable whenever two distances can end in a common point. The method of halving is simpler to administer since it is a bisection with one stimulus omitted, and the proof of its value will lie in its successful use. It is starting out well. The method of fractionation has been used even less, and I should not expect it to give as constant results as the other methods. The ultimate judgment in scientific measurement is the judgment of *equal* or *not-different*. In the methods of bisection and halving, the observer is really making an equation between the two halves. In the method of fractionation, if the ratio required be one-third, the observer has to judge that the given third is half of the remaining two-thirds, an involved judgment. One has only to remember what it is like to estimate two-sevenths of a specified distance to see how complicated with imaginal points fractionation may become.

It becomes plain that there are available two general measures of sensation: the DL and the comparable but more accurate σ of the psychometric function, on the one hand, and the sense-distance on the other. Are the two fundamentally different or the same? In other words, do all DL—we speak here about DL since σ has not yet been widely used in this context—do all DL represent the same sense-distance? That is the present form of the eighty-year old problem: are all jnd equal? And the answer to it is: No, not necessarily. The evidence is new and it comes from the field of hearing. The DL for pitch seem to represent approximately equal sense-distances,[13] but the DL for loudness do not.[14] When the loudness of a tone, as determined by the method of halving, is plotted against the number of DL that the tone lies above its absolute threshold, the function is not linear. The sense-distance that corresponds to the DL increases with the loudness. In general it increases most for low tones and less for high tones. The corresponding data for visual lightness have not yet been determined, although Judd makes it appear that DL for the saturation of yellow represent about equal sense-distances.[15]

There is a good deal of resistance to accepting this conclusion about the DL because it seems to be contrary to common sense. It is said that jnd must all be equal because they are all just noticeable. As a matter of fact, in that sense, they are equal: they are determined by identical operations and must be equal. For the same reason all DL must be equal in the sense that they represent the same frequencies of judgments. Fifty per cent must equal fifty per cent. The important point,

---

[12] The method has been used by P. H. Geiger and F. A. Firestone, The estimation of fractional loudness, J. *Acous. Soc. Amer.*, 5, 1933, 25-30, and also in this symposium by S. M. Newhall, The ratio method in the review of the Munsell colors, *infra.*

[13] Stevens and Davis, *op. cit.*, 84-89, 95-97.

[14] Stevens and Davis, *op. cit.*, 147-152.

[15] Judd, *loc. cit.*

however, is that, although all DL may be thus operationally equivalent, they are not necessarily equal *as sense-distances,* for the operations used in their measurement as sense-distances are different from the operations of their measurement as frequencies. The fact is that the frequency with which a difference is perceived is no guarantee of the exact size of the difference that is perceived. Similarly there is no particular reason to suppose that all jnd are equal. The jnd of the instant occurs under a release mechanism; something breaks up the levelling tendency and the two sensations jump apart, so that suddenly, with continuous change of the stimulus, a difference is perceived. The jump, however, may be large or small. All the evidence

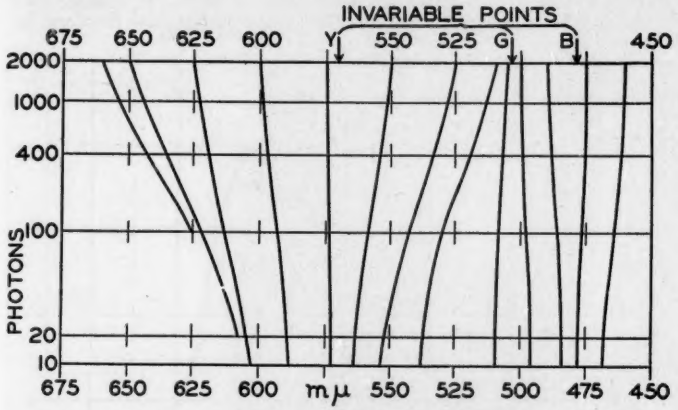


FIG. 6. CONTOURS FOR CONSTANT HUE ON THE SPECTRAL DIAGRAM: THE BEZOLD-BRÜCKE PHENOMENON
(From D. M. Purdy, this JOURNAL, 49, 1937, 314.)

now points to the fact that that distance which is jnd is not necessarily always the same size when it is measured as a sense-distance.

A diagram for sense-distances. Scientists interested in color ought to think more about a diagram for sense-distances. Even if no one ever had the time to determine the 10,000 values that might make it possible to construct such a diagram for monochromatic lights, still contemplation of the problem would impress investigators with the need for more data of the sort that would be especially useful in the measurement of tolerances. May I describe, for this purpose, the diagram that I should like to see constructed?

We should have first, I suggest, what might be called a spectrogram for the *lightness* of monochromatic lights. (The term *spectrogram* is analogous to *audiogram* in researches on hearing.) It would be a graph plotted against wave-length and energy for constant lightness, a family of curves with lightness varying as the parameter. Such graphs are the plots of the basic functions of which the visibility curves are the reciprocals, and we have had such data ever since König. I want, however, to have the different curves of the family separated by equal sense-distances of lightness. Thus this spectrogram would consist of a batch of U-shaped curves, which measure lightness as well as wave-length and energy. In other words,

the total graph would be really a representation of a surface plotted against three dimensions—wave-length, energy and lightness—and giving the interrelation of these three variables for pure light. It would be a trough with its longitudinal axis perpendicular to the plane of the axes for wave-length and energy.

Next I want the corresponding family of curves—that is to say, the corresponding surface—for wave-length, energy and *hue*. This is the surface that gives the complete summary of the Bezold-Brücke phenomenon, and Fig. 6, giving Purdy's data for this phenomenon, shows what the surface might be like.[26] It happens that the ordinate here is scaled in photometric instead of radiometric units, but that does not matter. Each curve shows how, when energy is changed, wave-length must also be changed if hue is to remain constant. The diagram gives these functions for 13

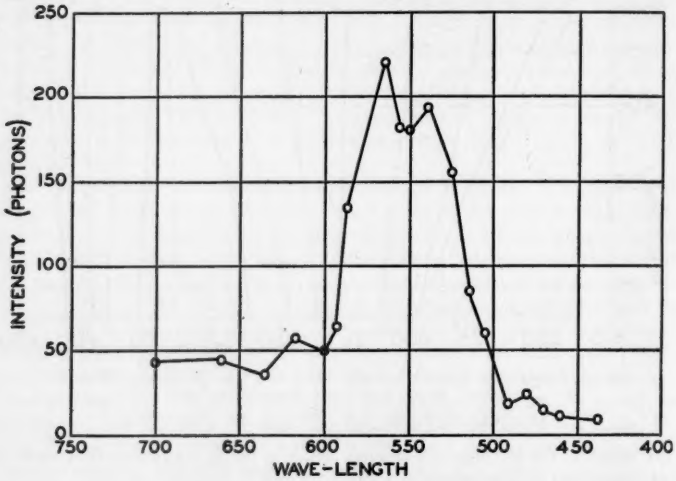


FIG. 7. PLAN VIEW OF CREST OF MAXIMAL SATURATIONS ON THE SPECTRAL DIAGRAM
(From D. M. Purdy, *Brit. J. Psychol.*, 21, 1931, 303.)

hues, but it makes no assumptions at all about the sense-distances between these hues. If the separations of the hues were known, then hue could be plotted perpendicular to this plane and we should have a surface that is slightly ogival, bending one way in the middle and the other way at the two ends.

Next I want a family of curves, and thus a surface, for the relation of wave-length, energy and *saturation*. Maximal saturation occurs at middle energies: very light and very dark colors are poorly saturated. Fig. 7 shows Purdy's determination of the energies which, for different wave-lengths, give maximal saturation.[27] One

[26] D. M. Purdy, The Bezold-Brücke phenomenon and contours for constant hue, this JOURNAL, 49, 1937, 313-315.
[27] Purdy, On the saturations and chromatic thresholds of spectral colours, *Brit. J. Psychol.*, 21, 1931, 283-313, esp. 302-304.

can think of this curve as the plan view of the crest of a ridge of saturation, in a diagram in which saturation is plotted vertically on a base for which wave-length and energy are the axes. This ridge, however, varies in height. Its profile would be given by the measurement of the different maximal saturations. It might look like Fig. 8, for here is the curve of Jones and Lowry measuring spectral saturation by the counting of DL's.[18] I have had to draw the curve down at each end, because it seems plain that the spectrum loses in saturation as well as lightness as it passes into invisibility. There are no well-saturated very dark colors.

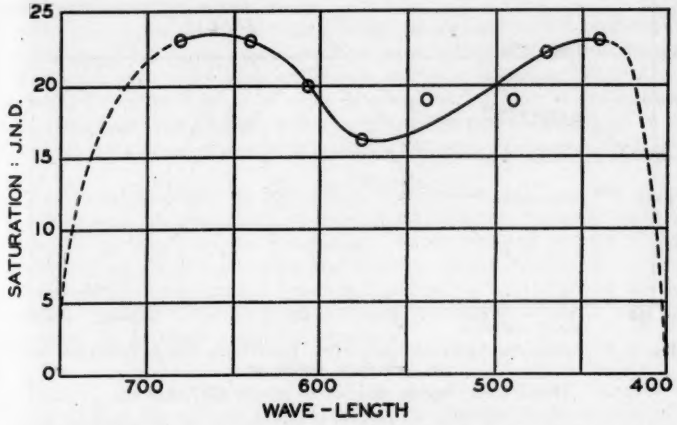


FIG. 8. ELEVATION OF CREST OF MAXIMAL SATURATIONS ON THE SPECTRAL DIAGRAM

(The data are from L. A. Jones and E. M. Lowry, *J. Opt. Soc. Amer.*, 13, 1926, 25-34, but the dotted extensions are the present author's idea.)

Now the lines of constant saturation would be the contour lines for this ridge. Fig. 9 shows one such contour.[19] The dotted line is the crest of the ridge—Purdy's data. The data of Jones and Lowry have been used to show how far down the ridge a contour of constant saturation-level might lie. Actually we want a set of these contours spaced at equal distances of saturation. Since the ridge has two peaks, the highest contours would be two separated rings. It is not, however, possible to draw them because we do not have the necessary measurements.

These three sets of facts together represent a surface in five dimensions. It is the complete spectrogram for monochromatic light. Its dimensions are wave-length, energy, hue, lightness and saturation. Since the figure is a surface, any two of the coordinate values determine the other three, once the shape and position of the surface are fixed.

---

[18] L. A. Jones and E. M. Lowry, Retinal sensibility to saturation differences, *J. Opt. Soc. Amer.*, 13, 1926, 25-34.

[19] E. G. Boring, Isochromatic contours, this JOURNAL, 49, 1937, 130-134.

It is in some such diagram that color tolerances might best be plotted. They would appear as little closed cells surrounding a point. They would not necessarily be five-dimensional hyperspheres, because the diagram would not assert any equivalence between a difference in hue and a difference in lightness. One might, however, hope—and this is my sole excuse for mentioning this elaborate figure—to get constant diameters for the cells parallel to the axes for hue, lightness and satura-

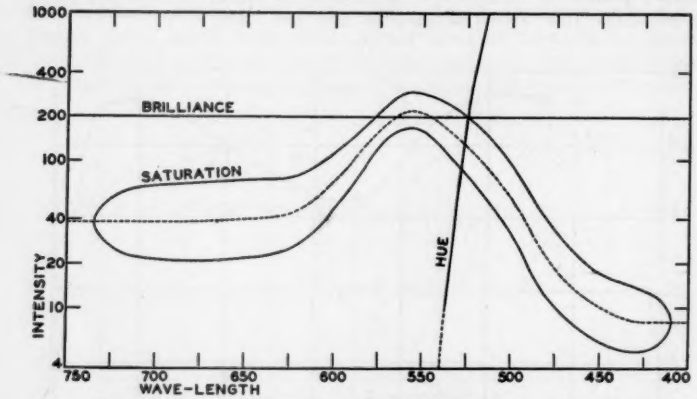


FIG. 9. ISOCHROMATIC CONTOURS FOR HUE, LIGHTNESS AND SATURATION ON THE SPECTRAL DIAGRAM
(From E. G. Boring, this JOURNAL, 49, 1937, 132.)

tion, since they are perceptual variables; that is to say, one might hope that, for a given purpose, tolerance for hue would be linearly related to sense-distance for hue, and similarly in respect of the tolerances for lightness and saturation. The measures of tolerance would not then show a linear relation to wave-length and energy, since these dimensions of the stimulus are not linearly related to sense-distance. In other words, perceptual scales would be found best for measuring perceptual differences.

THE RATIO METHOD IN THE REVIEW OF THE MUNSELL COLORS

By SIDNEY M. NEWHALL,* Johns Hopkins University

There are several reasons why the review of the spacing of the Munsell colors is of interest in connection with the special subject of color tolerances. The ratio method,[1] which has been relied upon from the beginning in this review, not only seems suitable for color tolerance determinations, but is actually being used for

* Chairman of the Subcommittee of the Colorimetry Committee on the review of the spacing of the Munsell Colors.

[1] D. B. Judd and D. Nickerson, Review of the spacing of the Munsell colors, *Mimeo. Statement,* (April) 1936, 1-6.

that purpose by Judd[3] and by Balinkin.[4] The present study provides the test of further use under complex conditions for this new method of color estimation.

The perfection of the Munsell spacing would provide a suitable system for the psychological expression of color tolerances. There is a need for such a system, and belief in its eventual availability has already lead to the development by Nickerson of a tolerance formula for use with it.[5] Incidentally, whether or not technical interest lies in tolerance problems, an adequate direct color comparator system would have the advantages of great facility and wide application.

This paper describes the ratio method and its application to the spacing problem and presents some preliminary results.

The ratio method. The ratio method is psychological and implies psychological measurement, but the general question of the possibility of color measurement need not detain us now. Color sensations may be unitary or indivisible, but they have magnitude; sensations can not be added or multiplied, but numerical values representing the magnitudes of their attributes can. The ratio method is employed in the important operation of estimating or assigning the values to represent the magnitudes.

This method apparently originates with L. F. Richardson, and it was first applied by him to color.[6] As generalized, it consists in the estimation by direct impression of the ratio of supraliminal sense magnitudes. The one magnitude is given or taken as standard, and the ratio of the other to it is then estimated directly.

There are really two forms of this method, both of which are applicable to color problems. The first form may be referred to simply as the ratio or R form and represented by the simple operational equation

$$(C_2/C_1)_A = R_A$$

in which C_1 represents the chosen standard or comparison color; C_2, the color to be evaluated; R, the estimate of the ratio; and the subscript A, the attribute with respect to which the estimate is made. Thus, the given color C_2 is judged equal to RC_1 with respect to A. Suppose, for instance, that C_1 and C_2 represent two patches of color and the observer's task is to estimate their relative lightness. He compares the two colors directly with each other and reports a quantitative judgment concerning the fraction or multiple which the lightness of C_2 is of C_1. In case of surface colors, A may become either H, S, or L, according to whether hue, saturation, or lightness is the attribute being estimated.

The second, or difference ratio method, is designated by R′ and represented by the analogous equation

$$(|C_1 - C_4|/|C_3 - C_1|)_{A'} = R'_A$$

in which $|C_2 - C_1|$ is the perceived standard difference or interval, $|C_4 - C_3|$ is the difference to be evaluated, and R′ is the estimate of the ratio of intervals with respect to A, the particular aspect under estimate.

---

[3] D. B. Judd, Specification of color tolerances at the National Bureau of Standards, *infra*, 418-428.

[4] I. A. Balinkin, Industrial color tolerances, *infra*, 428-448.

[5] D. Nickerson, The specification of color tolerances, *Textile Research*, 6, 1936, 505-514.

[6] L. F. Richardson, Quantitative mental estimates of light and colour, *Brit. J. Psychol.*, 20, 1929, 27-37.

The observer records his estimate either by linear extents or by numbers which seem to him to be in the ratio of the sense magnitudes in question. Perhaps the easiest method is to adjust the relative positions of the samples themselves until a space ratio represents the sense ratio. This method was demonstrated by throwing on the projection screen three color patches in a row, as C_4, C_2, C_1 in Fig. 1. These colors were in order with respect to a particular attribute in which they differed obviously. The intermediate color, C_2, was moved back and forth between the other two until a position was attained for which the space ratio, C_4 to C_2 : C_2 to C_1, seemed to equal the sense ratio, $C_4 - C_2 : C_2 - C_1$.

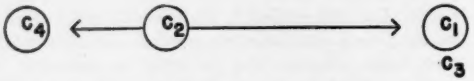


FIG. 1. ADJUSTING THE SPACE DISTANCES TO REPRESENT THE SENSE DISTANCES

A more generally practicable recording method is graphically to indicate rather than actually to make, the representative space adjustment. This is illustrated by Richardson whose subjects quantified a certain pink by making a check mark on a straight line representing a white-to-scarlet continuum.[6] It is also illustrated by the form of vector notation which is being used in the review of the spacing of the Munsell colors. Fig. 2 suggests the application of graphic recording to the R form. *O* represents the attributive magnitude of C_1 by a line of any convenient length. He then draws a second line, representing C_2, of such length relative to the first line that the ratio formed by the two lines is the same as the ratio of the two colors. Fig. 3 shows the same kind of recording in case of the R′ form of the ratio method. Here one line is drawn to represent the sense difference or sense distance

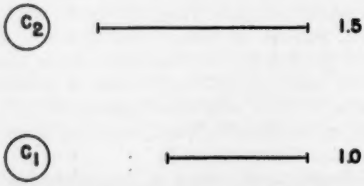


FIG. 2. MODE OF RECORDING ESTIMATES MADE BY THE SIMPLE RATIO METHOD, R

between C_3 and C_1, while the other line is so drawn that its length relative to the first will be the same as $C_4 - C_3$ relative to $C_3 - C_1$.

A third method, one which depends more on *O*'s unaided judgment, is to record a number indicating what the appropriate adjustment would be. Without bothering to draw any lines, *O* simply decides what multiple or fraction the one chromatic magnitude is of the other and assigns the number. If the methods are interchangeable, this number would be the same as the corresponding ratio of lines. In Fig. 2,

[6] L. F. Richardson, Quantitative mental estimates of light and colour, *Brit. J. Psychol.*, 20, 1929, 31-32.

for instance, the line representing C_4 is 1.5 times the line of unit length representing C_1; consequently a numeric estimate of about 1.5 is to be expected also.[7]

All three methods assume that, given the colors, a ratio is to be found. This is the obvious premise in our general problem of color evaluation. A fourth recording method would be required to cover the case where *O* starts with a given numerical or spatial ratio and has the task of painting, adjusting or otherwise producing a sample which will conform to it. This reverse problem is posed in the traditional sense distance procedures of the psychological laboratory.

The ratio method, being a sense distance method, is related to the four specific sense distance procedures defined by Boring.[8] One of these procedures, known as fractionation, is essentially the simple R form of the ratio method. The only difference seems to be that fractionation has been restricted by definition to the case where *O* adjusts a stimulus to produce a prescribed ratio.[9] A special case of fractionation is the halving procedure. The other two procedures, equation and

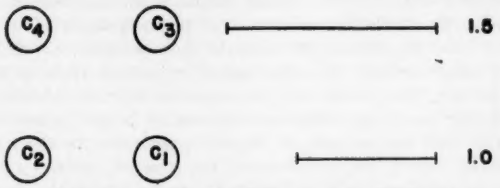


FIG. 3. MODE OF RECORDING ESTIMATES MADE BY THE DIFFERENCE RATIO METHOD, R′

bisection, resemble the R′ ratio form in that they deal in differences, but apart from the special case of equality, they do not involve the ratio principle at all.

The fact that estimates by both the R and R′ forms of the ratio method are recorded in the same manner should not conceal the fundamental difference in the forms themselves. The direct comparison of one sensory magnitude with another is a different operation from the comparison of intervals between sensations. Experiment and introspection evidence a different set which may, therefore, lead to different results. Even in the limiting case, equivalence cannot be assumed or even

---

[7] The same would hold for Fig. 3 because the recording methods are identical. Incidentally, Fig. 1 like Fig. 3, illustrates the R′ form and the two are simply related arithmetically. The situation in Fig. 1 differs in two respects from that in Fig. 3, *i.e.* (1) the ratio of estimate is $C_4 - C_2/C_2 - C_1$ instead of $C_4 - C_2/C_2 - C_1$, and (2) the attributive equality of C_1 and C_3 is assumed. Under these conditions it is easy to show that $(C_4 - C_2/C_2 - C_1) + 1 = C_4 - C_2/C_2 - C_1$. So, if an estimate has been recorded by adjusting the position of an intermediate color, the unit constant might be added to the result. Thus, in this instance, $C_4 - C_2/C_2 - C_1 = 0.5/1.0 = 0.5$, and $0.5 + 1 = 1.5$. This result is seen to agree arithmetically with the more usual R′ form, but whether psychological differences would permit agreement in actual practice is not known.

[8] E. G. Boring, The psychophysics of color tolerance, *supra*, 384-394.

[9] E. B. Newman, J. Volkmann, and S. S. Stevens, On the method of bisection and its relation to a loudness scale, this JOURNAL, 49, 1937, 134.

expected. The perception of a certain magnitude of saturation, for instance, is a different process from the perception of the interval between that magnitude and no saturation. Even if C_4 and C_1 in Fig. 3 both represent zero saturation and $[(C_4 - C_3)/(C_2 - C_1)]_S$ is equal arithmetically to $(C_4/C_3)_S$ it fails to be equal to $(C_4/C_3)_S$ introspectively. Lack of equality in reported results would raise a fundamental issue in color measurement.

Other interesting characteristics of the ratio method which have emerged or gained emphasis through use in experiment are, for convenience, presented here rather than in a subsequent section on results.

(1) Despite greater complexity, the R′ or difference form has a technical advantage over the R form because the former yields two points on the attributive scale of estimate and therefore defines it. Zero is defined by zero difference, that is, $|C_4 - C_3| = |C_2 - C_1|$, and a finite amount is represented by the unit amount or chosen standard, $|C_2 - C_1|$. In the case of the R method, the unit amount is also provided by the chosen standard, C_1, but the zero is unknown. Thus the R form is characterized by the same strange feature as the fractionation and halving procedures, *i.e.* the attributive estimate must be independent of a perceptually defined scale. That the estimate can be made is, nevertheless, a well known fact.

(2) Surprisingly enough, the ratio method appears to apply to the so-called qualitative variable, hue, as well as to saturation and lightness; at least this is true for the relatively small hue differences encountered in the present review. The reason may be that the hue scale is virtually quantitative for short ranges lying between 'unitary' colors. The evaluation of R_H, however, involves a limitation in addition to that mentioned in *(1)*. Though C_2 can be estimated as departing more or less from C_1, it seems impossible even to imagine a zero for the hue scale. Consequently *O* can not even imagine C_2 as being more or less 'hueful' than C_4. Apparently, the most he could do would be to estimate the ratio of the departures of C_2 and C_4 from the nearest 'unitary' color or from some other reference color. But that would really be an application of the R'_H form to which the above limitation does not apply.

(3) The work of Judd and of Balinkin indicates that the ratio method can be used in the evaluation of over-all color difference without explicit attributive abstraction. The importance of this discovery for visual psychophysics and the determination of psychological color tolerances is not generally recognized.

(4) Little is yet known regarding precision of estimate by the ratio method. The limited data available suggest that the precision is definitely lower than that of the standard psychophysical methods. In a study of the attribute of loudness, the uncertainty of the individual estimate is 0.6 of the mean.[10] In the study on the placement of pink—mentioned above, the standard deviation of the individual estimate is 0.3 of the mean.[11] Balinkin's research yields valuable information on this matter. Considerable interest naturally attaches to the order of precision under the complicated conditions of the Munsell review, the figures of which will be available eventually.

---

[10] L. F. Richardson and J. S. Ross, Loudness and telephone current, *J. Gen. Psychol.*, 3, 1930, 296, 304.

[11] L. F. Richardson, Quantitative mental estimates of light and colour, *Brit. J. Psychol.*, 20, 1929, 37.

The spacing problem. An important requirement in the construction of an ideal surface color solid is that equal sense distances shall correspond to equal linear extents. How closely this condition can be approximated in Euclidean space is still uncertain. A second requirement often imposed is that the dimensions of space shall represent the attributes of color, *i.e.* hue, saturation, and lightness, or in Munsell terms, hue, chroma, and value. The spacing problem consists in adjusting the locations in space of the regular Munsell samples until they are in essential agreement with the two requirements just mentioned. This amounts to the correction of imperfecions in a real chromatic skeleton of 400 units so that it will more properly fit into an ideal chromatic body.

The psychologically equi-stepped solid, which the Munsell system approximates in part, includes a vertical achromatic axis upon which is stacked a finite series of horizontal, iso-lightness planes ranging from black below to white above. Surrounding the axis is a series of concentric, iso-saturation cylinders ranging from minimum within to maximum without. Radiating from the axis is a closed series of vertical iso-hue planes arranged in the spectral order and including the purples. The 7 Munsell constant-value charts and the 20 constant-hue charts[12] correspond, respectively, to equi-stepped iso-lightness and iso-hue planes of the ideal figure. The saturation cylinders as well as the planes are broken and limited by the bounding surface of the solid.

Fig. 4 exhibits a Munsell solid as photographed from five different angles separated by 72°, to show the shape relative to the loci of the five principal Munsell hues, R, Y, G, B, P. The achromatic axis is indicated, and constant-value and constant-hue planes can be visualized as horizontal or vertical cuts through the solid. Great differences in horizontal distances from the axis to the surface indicate great differences in the chroma maxima for the various hues and lightnesses.

The Munsell solid approximates the psychological ideal in regard to its three dimensions taken separately but not collectively. This is because the scaling units of hue, chroma, and value are far from equivalent perceptually. Fig. 4 shows the solid to be half as high as it is broad, but when the scaling units have been roughly equalized and the proportions of the solid more closely approximate the ideal, the height and breadth become about equal.

This picture may help to make the spacing problem seem concrete and to suggest that spacing adjustments in three dimensions are indicated. But it is well to remember that the particular shape of the Munsell solid is really a secondary matter, depending not only on the choice of units but also on the availability of sufficiently permanent pigments for practical purposes.

Though this solid is much smaller than the theoretical pigment maximum,[13] the contained number of just discriminable color elements is undoubtedly in the millions. In comparison, the 400 Munsell colors may seem a very small number, and the sensory distances among them must be very large under favorable conditions of observation. But adjacent Munsell samples do not seem very widely separated under ordinary conditions of observation and so, perhaps, do not form too bare a skeleton. This statement gains support from the allowance by Judd and Kelly, for drug

[12] Munsell Color Co., *Munsell Book of Color*, 1929.

[13] D. L. MacAdam, Maximum visual efficiency of colored materials, *J. Opt. Soc. Amer.*, 25, 1935, 361-367.

description, of about one color name per 20,000 colors.[14] In any event, if these colors can be correctly located, further refinement might follow or be proved not feasible.

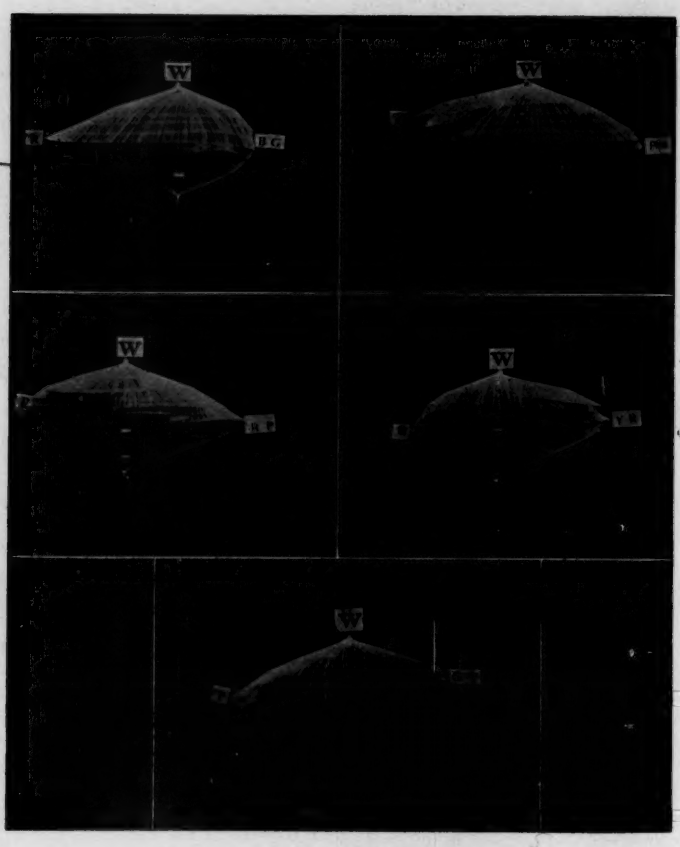


FIG. 4. THE ROUNDED MUNSELL SOLID

The question of the degree of accuracy requisite to the successful completion of the spacing problem is itself a tolerance problem. Several considerations may be mentioned: (a) Since the whole point of the respacing effort is to reduce discriminable irregularities, it would seem ideal to take as the measure of tolerance the differential threshold under favorable conditions of viewing. Probably this

[14] D. B. Judd and K. L. Kelly, Scientific color naming of drugs, *J. Amer. Pharm. Ass.*, 27, 1938, 208-211.

ideal may be approached but not reached with the present procedure. (b) There is the fact that the Munsell papers themselves exhibit a limited amount of perceptible variation in samples assigned the same notation. Thus the variability of the papers might be taken as the tolerance limit. (c) On the other hand, variability of the papers imposes no necessary limitation on accuracy. If that variability is found to be too high, as evaluated perceptually, it can be reduced by adjusting the manufacturing tolerance. The master series of samples deposited with the National Bureau of Standards by the Munsell Color Company can be taken as standard and held constant indefinitely by appropriate analysis and colorimetric specification.

The general attitude taken in the present review is that any substantial improvement in spacing which may prove possible in a few years' time will be well worthwhile, whether or not the ideal of the DL tolerance limit can be closely approximated. Indeed, the DL under favorable conditions is so small that considerable departures from it would be tolerable in many circumstances.

The ratio method applied to the spacing problem. The suitability of the ratio method for the spacing review may be briefly explained. First of all, the method is sufficiently flexible to permit the use of a relatively general or free instruction to the observer. In this new application of a new method the desirability of providing *O* with considerable latitude may be fairly evident; such a provision might lead to more representative results or to the discovery of better ways of observing than would a more rigid instruction. Another reason for having this free instruction is the great variation in the experience of the participating *Os*; the instruction needs to be relatively general to be comprehensible to all of the *Os*. Even so, the complete instructions cover nearly five typewritten pages.

The fact that the ratio method is rapid is an advantage in making the large number of color comparisons required. It is the time factor which makes the thorough-going application of presumably more precise methods seem impracticable. The experimental time factor imposed by the standard threshold methods would seem to prohibit their use.[35] Similarly, a computational time factor tends to exclude multidimensional psychophysics of the sort suggested by M. W. Richardson.[36]

The use of the ratio method seems peculiarly significant because, as Judd has pointed out, it amounts to the direct application of the definitions of the attributes of color. Unless the definitions are changed or other methods are proved comparable, only the ratio method is strictly applicable.

Observers are making the estimates of hue, value, and chroma on the Munsell constant-value and constant-hue charts. Both forms are being applied: the color ratio, or R form, consists here in the direct comparison of several samples with respect to the given attribute; and the difference ratio, or R′ form, calls for the comparison of the attributive differences between samples. In the one case *O* asks himself, "How much does each of these colors differ in this attribute in regard to the particular degree of it which I select as my comparison standard?" In the other case, "How much does the attributive interval between this pair of colors differ in regard to the degree of attributive difference which I have taken as standard?" In both cases the estimates of departures from standard are recorded in

[35] J. P. Guilford, *Psychometric Methods*, 1936, Pt. I.

[36] M. W. Richardson, Multidimensional psychophysics, *Psychol. Bull.*, 35, 1938, 659-660.

vector notation in a record form similar in shape to the color chart. The effort is to represent various degrees of departure by drawing vectors of proportionate lengths.

Hue is estimated from the constant-value charts. The series of samples in a given 'constant' hue radius is isolated by masking the remainder of the chart and examined for variations in hue. If such are observed, the effort is made to estimate a representative or average hue and indicate departures therefrom by tangential vector notation. The same procedure is to be followed in turn for each 'constant' hue radius on the chart. Next, the spacings of the hue radii considered as wholes are examined, and in any cases where hue differences seem out of proportion to the space differences, a tangential vector is affixed to the end of the radius to indicate rotation.

Chroma, like hue, is estimated from the constant-value charts. The samples in a given 'constant' chroma ring are masked off, and *O* looks them over in an effort to select a sample of representative chroma to serve as standard. Any sample which differs in chroma from the standard is then evaluated in terms of it, and a suitable radial vector is drawn to indicate the displacement. Subsequently, the masks are removed, and the 'constant' hue radii are examined for irregularities in the size of the chroma steps between adjacent colors. After selection of a representative step or chroma difference, each of the other steps can be evaluated in terms of it.

Chroma is estimated again, as a check, from the constant-hue charts. A constant chroma column is masked off, and the R form is applied. After each column has been observed, the masks are removed, and the equality of the chroma steps is examined by the R′ form.

Incidentally, most subjects seem to find the applications of the R′ form to chroma the most difficult part of the entire procedure. This is because of the considerable number of chroma steps to be taken into consideration.

Value is also estimated from the constant-hue charts. 'Constant' value rows are masked off and evaluated by the R form, after which the masks are removed, and the size of the value steps between successive 'constant' value rows are adjusted by the R′ procedure. In the case of value, a special difficulty in applying the vector recording technique led to the substitution of numerical estimates of the spacing in terms of the value step unit.

Since there are 7 constant-value charts and 20 constant-hue charts, and since each of these 27 charts is prepared with three backgrounds—white, gray, and black —there are to be examined a total of 81 charts. The task is long and so exacting that it can not be relegated to periods of fatigue. Those few who complete it usually take more than a year to do so.

Complications in the application of the ratio method. Psychophysical methods differ from physical in that none of the former seems satisfactory to the naïve *O*. The reason is that the critical judgments necessarily fall within a range of uncertainty and seem more or less like guessing. Incidentally, this condition is meritorious, for the consequent variability is more indicative of sensitivity than lack of precision. Some of the *O*s in this review find the judgmental situation so "difficult," "impossible," or "terrible" that there is special reason to consider the problems reported by them.

(1) There is the problem of complex anisochromatic comparisons forcing, as it does, the abstraction of the attribute to be estimated. Attributive abstraction is

peculiarly difficult in cases of strong association; for instance, when saturation is the judged attribute while hue is constant and value is variable, there is the strong tendency to estimate lower-value colors as of higher saturation. That is a relation often found in everyday life. Another familiar problem is that of equating for lightness colors of differing hue. A similar difficulty which is also emphasized in case of unsophisticated subjects is that of equating chromatic and achromatic colors in respect to lightness. The untrained *O* usually sees chromatic samples as lighter relative to achromatic samples, than does the trained *O*. Inspection of data from the constant hue charts reveals the interesting fact that the vectorial lightness adjustments for the chromatic samples are never downward. Wherever an adjustment is made, it indicates that the chromatic sample seemed too bright relative to the achromatic of the same Munsell value.

(2) *O*s making saturation estimates by the R′ method report greater facility in difference comparison when both differences are well about achromaticity than when one terminates there. The trouble seems to be that a first step extending from zero to some chroma appears as a qualitative step and so tends to give the impression of an infinite quantitative step. The second, step, or difference, from some chroma to more chroma is, on the other hand, a definitely limited quantitative step. Though a very low chroma level is actually attained in the first step, the step itself is relatively great. This 'zero difficulty' of the R′ method bothers only some of the *O*s.

Possibly related to it is a certain change of attitude correlated with a change of chroma level. When observing chroma differences near neutral, some observers report that they 'just look for any difference' whereas at the higher chroma levels they think more definitely in terms of chroma. Such a change of attitude might be expected to occur naturally on the basis of least effort. Looking for a mere difference would be easier when the detection of chroma is difficult, whereas looking for a chroma difference would be easier when chroma is present in obvious degree.

(3) The long recognized difficulty of spatial or temporal separation of samples is evidently present in this study where widely separated parts of a chart are compared. Of incidental interest are preliminary data indicating surprisingly good color matches when time differences between standard and match are over two minutes!

In the face of the difficulties, what do the *O*s do? All that can be done here is to note a few common tendencies apparently present. (a) The series of colors, or color differences, to be judged is first surveyed as a whole, and the *O* essays a decision with respect to rank order for the attribute in question. Often no progressive graduation is found; and instead of arriving at an order, *O* is able to identify a few extreme values and a number which seem about equal. (b) After some such preliminary survey, *O* may proceed to the assignment of vectors. A common practice is to draw in vectors representing the more extreme displacements first, and then, partly on the basis of those, to fill in the shorter vectors indicating the lesser irregularities. (c) There is the tendency to 'carry' a standard subjectively even though some particular perceptual standard has been chosen. *O* may occasionally refer to the standard sample for verification, but actually make various comparisons without it.

This 'absolute' procedure and the ranking procedure mentioned above were both hit upon without instruction, probably because they facilitate *O*'s difficult task. Certainly it is easier to decide which of two sensory magnitudes is greater than to

estimate how much greater; and it is sometimes easier when making numerous comparisons to carry a subjective standard than to refer back each time to a perceptual one.

Evidently, it would be a mistake to conclude that the ratio method has been applied to the spacing problem in anything like pure form. Neither, in advance of repetitive data, can one draw any definite conclusions with respect to the self-consistency of the given *O*. On the other hand, the instructions call for so thorough

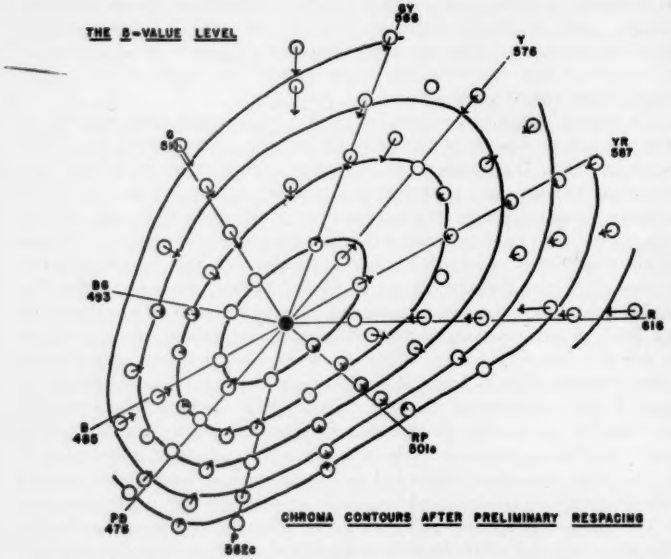


FIG. 5. ADJUSTED IRREGULARITIES AND CHROMA CONTOURS AT THE 5-VALUE LEVEL

an examination of the charts that marked irregularities can scarcely be overlooked.

Preliminary results. Most of the results now available have to do with the nature and technique of the ratio method; and these results have already been described *en passant*. For the purposes of this paper, the status of the respacing itself may be dismissed in a few words.

Preliminary results on the respacing of saturation were published by Nickerson last summer.[17] The Munsell colors of each constant-value chart were plotted in ICI trilinear coördinates, utilizing the Hardy-Glenn-Killian data. Then a summary was made of the chroma estimates for each individual Munsell color. Corresponding to the summary estimate, a radial arrow was attached to each point in the diagrams which was not confirmed by the consensus. The direction of this vector arrow would

[17] D. Nickerson, Use of ICI tri-stimulus values in disk colorimetry, *U. S. Dept. Agriculture*, 1938, 1-17.

be inward or outward according to the consensus of estimates, and its length would be long or short according to whether the sample was judged to be considerably or only slightly out of line. After all the radial arrows had been marked in, representative chroma contours were established by drawing smooth curves through the scatters of the arrow-heads. Smoother and more sweeping curves could be passed through the arrow-heads than through the original unadjusted points from which the arrows were drawn. This result seemed to indicate that the visual estimates had reduced some of the irregularities in chroma, and, in general, that the ratio method was proving useful.

This result is illustrated by Fig. 5, which shows the plot of the Hardy-Glenn-Killian determinations for the 5-value Munsell chart, together with Nickerson's smoothed curves and arrows for the gray ground. It suggests that many minor adjustments will probably prove desirable. The figure also shows that the ICI coordinate system is far from an equal chromaticity system and so not an ideal system of reference. Still, the departures from equal chromaticity do not seem sufficiently sudden to give rise to any serious misinterpretations, and it is probably best to utilize this standard system.

Preliminary comparison of results secured from use of the black, white, and gray charts indicates that background reflectance plays a significant rôle. This was to be expected from the results of Helson[18] and Judd.[19] The stimulus-specifications for the ideal solid vary markedly with the background, especially at the extreme values. Thus the solution of the spacing problem is not simple.

Summary

(1) The ratio method consists in the estimation by direct impression of the ratio of supraliminal sense magnitudes. The one magnitude or interval is taken as standard, and the ratio of the other to it is then estimated.

(2) The spacing problem is to detect and correct errors of allocation in surface color space of the 400 regular Munsell colors. The readjustments are being made for the samples as mounted on white, on gray, and on black (cardboard) grounds.

(3) The application of the ratio method to the spacing problem proves to be a complicated, meticulous, and lengthy process of choosing vectorial equivalents for sensory magnitudes. The principal attributes are evaluated separately under a flexible instruction.

(4) Os are beset by the real and technical difficulties of the first chroma step, attributive abstraction, and separation of standard and sample. They do not use the ratio method in pure form, but resort to the methods of ranking and single stimuli in the effort to achieve results; and results are achieved.

(5) Preliminary results on the respacing of saturation tend to confirm the usefulness of the method and to suggest that many minor adjustments may prove desirable. Present data indicate that adjustments also will be necessary if the colors are to be corrected for gross variations in background reflectance.

---

[18] H. Helson, Fundamental problems in color vision. I. The principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination, *J. Exper. Psychol.*, 23, 1938, 439-476.

[19] D. B. Judd, Hue, saturation and lightness of surface colors with chromatic illumination, *Bull. Amer. Phys. Soc.*, 14, 1939, 23.

COLOR TOLERANCES AS AFFECTED BY CHANGES IN COMPOSITION AND INTENSITY OF ILLUMINATION AND REFLECTANCE OF BACKGROUND

By HARRY HELSON, Bryn Mawr College

The hue, lightness, and saturation of any object depend upon a complex of conditions, chief of which are spectral reflectance of sample and background, spectral energy distribution of illuminant, and state of the visual mechanism as determined by its properties as a receptor and mode of functioning. Since color tolerance concerns the extent to which colors match, or if they do not match, in what respects they differ, any factors which affect the color of standard and variables have importance for the problem of tolerance. How illuminant, background, contrast, adaptation, and so-called constancy operate in literally 'coloring' objects seems to be nearing solution after recent work with strongly chromatic illuminants and backgrounds of high, medium, and low reflectance.[1] The principles governing color changes which objects undergo in strongly chromatic illuminations throw light also

TABLE I

REPORTS OF 16 Os ON HUE OF MUNSELL AND ACHROMATIC SAMPLES UNDER PALO DAYLIGHT ILLUMINATION

	S	R	yR	RY	rY	Y	gY	YG	yG	G	bG	BG	gB	B	rB	RB	bR	A
R	3/7	12	3														1	
R	5/10	10	5														1	
R	7/6	13	2														1	
YR	4/5		1	4	9	2												
Y	5/7					10	6											
Y	7/8					15	1											
G	3/4								2	11	3							
G	5/7								3	13								
G	7/7								4	11	1							
B	3/5												6	10				
B	5/6												7	9				
B	7/4											1	1	14				
P	3/6														6	5	5	
P	5/6															10	6	
PB	7/4													16				
A	10/0																	16
A	5/0																	16
A	0/0																	16

on phenomena encountered in ordinary viewing situations. Hence consideration of the facts discovered under 'abnormal' conditions of viewing is of value in bringing into sharp relief the principles operating in all visual functioning.

Before discussing factors affecting color tolerances more or less common to all

[1] D. B. Judd, Surface color, *J. Opt. Soc. Amer.*, 25, 1935, 44; H. Helson, Tridimensional analysis and the non-film modes of color appearance, *ibid.*, 27, 1937, 59; Fundamental problems in color vision, I. The principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination, *J. Exper. Psychol.*, 23, 1938, 439-476.

Os, let us consider briefly individual differences in color vision,[3] due to a variety of causes, which preclude absolute agreement on color naming and color matching no matter how favorable the viewing conditions may be. Data on surface colors show that some *Os* do not tolerate a sample as a representative of a hue which others give it. The results in Table I, wherein reports of 16 *Os* are given for 15 Munsell samples and three achromatic samples of less aristocratic origin under Palo daylight illumination, show that none of the samples supposed to exemplify one of the four unitary hues, Red, Green, Yellow, Blue, was given a single hue name by all *Os*. In every case some *Os* report a second, minor component in the unitary hues. Thus the three Munsell reds, R 3/7, R 5/10, R 7/6, are reported as having a *yellowish* or a *bluish* minor component by some *Os*. Similarly more than one hue-name characterizes the reports by some *Os* on the other 'unitary' hues.

While a large majority of the *Os* report the same hue, dissenters may be put into two groups, one reporting as a secondary component a hue on one side of the principal one (*yellow* in G ¾), the other reporting as secondary component a hue on the other side of the principal one (*blue* in G ¾). On the other hand, some binary hues, for example YR ⅚, are seen as unitary by some *Os*. Agreement is much better, as Table I shows, with respect to binary hues than unitary hues. It is as if the unitary hues either do not have sufficiently *strong* secondary components to be visible to all *Os*, or sufficiently *weak* components to remain invisible so far as all *Os* are concerned. Practically, we may put it thus: a merchant selling blue shirts would find from data such as these what percentage of his customers would not tolerate his goods but would object that they are off-color, say greenish or reddish blue instead of a good blue. Furthermore he would be more apt to find agreement among his customers on binary hues than on so-called pure or unitary hues.

Apart from individual differences such as we have just discussed two of the chief factors influencing color tolerances are the composition of the illuminant in which the samples are viewed and the reflectance of the background on which they are placed. A clear-cut decision as to the rôle of each of these factors could only be made after many samples were observed in a variety of strongly chromatic illuminations on backgrounds of different reflectances because illuminant effects depend largely upon background influences.

Inspection of Table II brings out clearly the combined effect of illuminant and background on color of objects. Taking the 18 samples as a whole we find that on black background the hue of that illuminant predominates while on white background the complementary hue is strongly in evidence as Table II clearly shows. With gray backgrounds both hues appear about equally often. For each background reflectance with any given set of samples there is established a certain level of adaptation identified by a reflectance above which appear illuminant colors and below which

[3] One of the chief sources of differences in reports on hue names, aside from practice and training, lies in the macular pigment which varies from person to person. Cf. W. O'D. Pierce, Individual differences in normal colour vision, Reports of the committee upon the physiology of vision, *Med. Res. Council, His Majesty's Stationery Office, London*, 1933, No. 181, 34. "The conclusion has thus been reached," says Pierce, "that the main source of variation among normal trichromats is due to the macular pigment varying greatly in density and to a less extent in dominant hue." We do not, however, imply this is the only factor governing the reports of our observers.

appear the complementary colors. Hence background furnishes a clue not only to hue but to saturation as well because samples near adaptation reflectance are either achromatic or low in saturation, highest saturations being found in either the lightest or darkest samples.[8]

TABLE II

HUE (H), LIGHTNESS (L), SATURATION (S) AND APPARENT REFLECTANCE (A), OF 15 MUNSELL SAMPLES AND THREE ACHROMATIC SAMPLES (S) IN FOUR ILLUMINANTS ON THREE BACKGROUNDS

			Red Illumination					
			White		Gray		Black	
	S	A	H	L/S	H	L/S	H	L/S
A	10/0	.800	R	9.0/4.0	RY	8.0/7.5	yR	8.0/9.0
R	7/6	.781	yR	9.0/3.0	RY	8.0/7.0	yR	7.5/9.0
R	5/10	.706	yR	8.0/3.5	yR	8.0/7.0	yR	7.5/9.0
Y	7/8	.577	RY	8.0/3.0	yR	7.5/6.0	yR	6.5/9.0
pB	7/4	.428	yR	7.0/2.5	RY	7.0/4.5	R	6.0/8.0
B	7/4	.381	rY	6.0/1.0	rY	6.0/5.0	R	5.5/8.0
P	5/6	.359	yR	6.0/2.0	RY	6.0/5.0	R	5.5/8.0
G	7/7	.329	R	5.0/1.0	rY	6.0/4.0	yR	5.0/7.0
YR	4/5	.297	yR	4.5/1.0	RY	5.0/4.0	yR	4.5/7.0
R	3/7	.276	R	4.5/1.0	RY	5.0/4.0	R	4.5/7.0
Y	5/7	.272	A	4.5/0.0	rY	5.0/4.0	R	4.5/7.0
A	5/0	.240	R	4.0/0.4	yR	5.0/1.5	R	4.5/3.5
B	5/6	.161	gB	3.0/1.0	YG	3.0/0.5	R	3.0/5.0
P	3/6	.153	gB	3.5/1.0	Y	3.0/0.7	bR	3.0/5.0
G	5/7	.087	GB	1.0/4.0	gB	2.0/2.0	yR	2.0/2.0
G	3/4	.072	GB	1.0/7.0	GB	2.0/2.0	yR	1.5/2.0
B	3/5	.064	bG	2.0/7.0	BG	1.5/2.0	yR	1.5/2.0
A	0/0	.030	gB	2.5/9.0	G	2.5/9.0	A	0.0/0.0
			Green Illumination					
			White		Gray		Black	
	S	A	H	L/S	H	L/S	H	L/S
A	10/0	.800	G	8.5/3.0	gY	9.0/9.0	gY	8.5/7.0
G	7/7	.596	yG	8.0/3.0	YG	8.0/5.0	yG	7.5/6.0
B	7/4	.581	yG	8.0/3.0	yG	8.0/5.0	yG	7.5/6.0
pB	7/4	.485	G	7.0/3.0	yG	6.5/6.0	yG	7.0/6.0
Y	7/8	.455	rY	6.5/4.5	Y	6.5/5.0	gY	7.0/4.5
G	5/7	.377	yG	5.0/2.5	yG	6.0/7.0	yG	6.0/7.0
R	7/6	.357	Y	5.5/2.5	Y	6.0/3.0	YG	6.0/4.0
B	5/6	.324	bG	5.0/3.0	bG	5.0/7.5	G	5.0/7.0
A	5/0	.240	A	5.0/0.0	gY	5.0/2.0	yG	5.0/2.0
P	5/6	.235	rB	5.0/2.0	YG	5.0/0.5	yG	5.0/5.0
Y	5/7	.221	RY	5.0/4.0	rY	5.0/4.0	yG	5.0/3.0
B	3/5	.133	rB	3.0/4.0	B	3.0/1.0	G	4.0/7.0
R	5/10	.110	bR	3.0/5.0	bR	4.0/4.0	gY	4.0/2.0
YR	4/5	.110	bR	3.0/4.5	yR	4.0/4.0	Y	4.0/2.0
G	3/4	.107	rB	3.5/4.5	rB	4.0/2.0	G	4.0/2.0
P	3/6	.097	rB	3.0/4.5	rB	3.0/3.0	G	3.0/2.0
R	3/7	.049	RB	1.5/7.0	bR	2.5/7.5	RY	2.0/1.5
A	0/0	.030	RB	2.0/5.0	RB	2.0/9.0	A	0.0/0.0

[8] For a more thorough discussion of adaptation reflection and its relation to saturation and lightness, see H. Helson, *ibid.*

TABLE II—*Continued*

Yellow Illumination

			White		Gray		Black	
	S	A	H	L/S	H	L/S	H	L/S
A	10/0	.800	Y	9.0/4.5	Y	8.5/9.0	Y	9.0/8.0
R	7/6	.618	RY	8.0/5.0	rY	7.5/7.0	rY	8.0/7.5
Y	7/8	.588	gY	8.0/4.0	Y	8.0/7.0	Y	8.0/7.0
B	7/4	.445	yG	7.0/5.0	YG	8.0/5.5	YG	7.0/5.5
G	7/7	.428	yG	7.0/7.0	yG	7.5/7.0	yG	7.0/5.0
pB	7/4	.405	gY	7.0/3.0	gY	5.0/4.5	gY	7.0/5.0
R	5/10	.404	yR	7.0/8.0	yR	7.0/9.0	RY	7.0/8.5
Y	5/7	.288	rY	5.0/1.0	gY	5.0/3.5	yG	6.0/5.0
P	5/6	.273	bR	5.0/3.5	RY	5.0/3.5	rY	6.0/4.0
YR	4/5	.246	bR	4.5/3.5	rY	5.0/3.5	rY	5.5/4.0
A	5/0	.240	gY	4.5/0.5	Y	5.0/1.5	Y	6.0/2.0
B	5/6	.185	bG	3.0/0.5	yG	3.0/3.0	yG	5.0/4.0
G	5/7	.171	bG	4.0/8.0	G	5.0/9.0	G	5.0/8.5
R	3/7	.151	bR	4.5/9.0	bR	5.0/8.5	yR	5.0/7.0
P	3/6	.119	bR	3.0/5.0	yR	3.0/4.0	RY	4.5/4.0
G	3/4	.078	rB	2.0/2.5	rB	2.0/2.0	G	3.0/2.0
B	3/5	.072	rB	2.0/2.5	rB	2.0/2.0	yG	3.0/3.0
A	0/0	.030	rB	1.0/2.0	RB	1.0/3.0	A	0.0/0.0

Blue Illumination

			White		Gray		Black	
	S	A	H	L/S	H	L/S	H	L/S
A	10/0	.800	B	9.0/3.5	B	9.0/6.0	B	8.5/7.5
B	7/4	.642	B	8.0/6.0	B	8.5/6.0	B	8.0/6.0
pB	7/4	.600	B	8.0/5.0	B	8.5/7.0	B	8.0/6.0
G	7/7	.493	gB	7.0/4.5	B	7.0/6.0	B	7.5/5.0
B	5/6	.472	gB	7.0/3.5	B	7.5/7.0	B	7.5/8.0
R	7/6	.346	rB	6.0/6.0	RB	6.0/5.0	rB	6.5/6.0
P	5/6	.310	rB	6.0/6.0	rB	6.5/7.0	rB	6.0/8.0
G	5/7	.267	G	6.0/7.0	G	6.0/7.0	BG	5.5/6.5
A	5/0	.240	A	5.0/0.0	A	5.5/0.0	B	5.0/5.0
B	3/5	.217	B	5.0/0.5	BG	5.0/1.0	B	5.0/6.5
Y	7/8	.159	rY	3.0/3.5	rY	4.0/3.0	RB	4.0/1.5
P	3/6	.153	RB	5.0/6.5	rB	4.0/7.5	RB	4.0/8.0
R	5/10	.110	bR	3.0/1.5	bR	3.5/3.0	RB	3.0/6.0
Y	5/7	.097	yR	3.0/1.5	rY	3.0/2.0	bR	3.0/2.0
G	3/4	.093	Y	3.0/0.5	yG	3.0/1.5	B	3.0/1.5
YR	4/5	.069	RY	2.0/1.0	yR	3.0/3.0	bR	2.0/2.5
R	3/7	.048	yR	2.0/1.5	R	2.0/2.0	bR	2.0/3.0
A	0/0	.030	yR	1.0/0.5	RY	1.0/1.0	A	0.0/0.0

Under ordinary conditions of viewing the illuminant is not strongly chromatic and hence these effects are not so patent but the principle just formulated operates in all illuminants. Illumination having only a tinge of chroma colors objects in the ways stated even though to a slight extent.[4] Mixture of chromatic (filtered) light with 2848°K in varying degrees made possible illuminants of varying chromaticities

[4] D. Nickerson (Artificial daylighting for color grading of agricultural products, *J. Opt. Soc. Amer.*, 29, 1939, 1-9) has reported that illuminant 'C' gave a 'too-yellow' effect in grading cotton which was remedied by increasing the color temperature of the illuminant from 6500°K to 7400°K.

and the results in Table III show that the hue of samples is a resultant of the 'white' light hue and the dominant chromatic component of the illuminant. Non-selective samples betray the chromatic component in the illuminant to an even greater degree than selective samples. The larger number of reports of illuminant

TABLE III

RETURN OF HUES TO DAYLIGHT APPEARANCE WITH ADMIXTURE OF 0.3 F.C. AND 4.5 F.C. 2848°K WITH 3.7 F.C. RED

3.7 f.c. R plus 0.3 f.c. 2848°K

	S	A	Wh		Gr		Bl	
A	10/0	.800	R	9.0/0.5	R	8.0/3.0	R	7.0/7.5
R	7/6	.759	RY	8.0/4.5	RY	7.5/7.5	yR	7.0/8.5
R	5/10	.671	yR	7.0/6.0	yR	7.0/7.0	yR	7.0/9.0
Y	7/8	.573	Y	7.0/4.0	rY	6.5/6.0	yR	7.0/8.0
pB	7/4	.429	rB	6.0/1.0	RB	7.5/4.0	bR	6.5/6.0
B	7/4	.391	gB	7.0/1.5	B	7.5/1.0	bR	7.0/5.0
P	5/6	.351	RB	6.0/2.5	bR	5.0/5.0	yR	7.0/7.0
G	7/7	.343	bG	6.0/4.0	bG	6.0/5.0	bR	6.5/5.0
YR	4/5	.288	rY	5.0/2.0	rY	5.0/5.0	R	6.0/7.0
Y	5/7	.271	G	5.0/0.5	gY	5.0/3.0	yR	5.0/6.0
R	3/7	.263	bR	5.0/3.0	yR	5.0/6.5	R	5.5/7.0
A	5/0	.240	gB	4.0/0.5	R	4.5/1.0	bR	5.0/6.0
B	5/6	.168	gB	3.5/3.0	B	5.0/3.0	RB	5.0/4.0
P	3/6	.150	rB	2.5/2.0	bR	3.5/4.0	bR	4.0/6.0
G	5/7	.100	bG	3.5/7.0	G	4.0/8.0	bG	5.0/5.5
G	3/4	.074	bG	2.0/3.0	bG	2.0/1.0	bR	4.0/2.0
B	3/5	.067	gB	2.0/4.0	gB	2.0/4.0	rB	3.0/3.0
A	0/0	.030	BG	0.5/3.0	GB	1.0/2.0	A	0.0/0.0

3.7 f.c. R plus 4.5 f.c. 2848°K

	S	A	Wh		Gr		Bl	
A	10/0	.800	RY	9.5/0.5	R	9.5/0.5	R	9.0/1.0
R	7/6	.622	yR	8.0/6.0	R	7.5/4.5	yR	7.5/5.0
Y	7/8	.548	gY	7.0/6.0	Y	7.0/6.0	Y	6.5/7.0
R	5/10	.458	yR	7.0/9.0	yR	7.5/9.0	yR	6.0/9.0
B	7/4	.455	gB	7.0/6.5	gB	7.5/4.0	B	8.0/2.5
pB	7/4	.436	rB	7.5/5.0	rB	7.5/3.0	rB	8.0/2.0
G	7/7	.429	G	7.5/7.0	G	6.5/5.5	GB	7.5/4.0
P	5/6	.301	RB	6.0/6.0	bR	6.0/6.0	bR	6.0/5.0
Y	5/7	.264	YG	5.0/5.0	gY	5.0/5.0	Y	5.0/6.0
A	5/0	.240	A	4.5/0.0	A	5.5/0.0	bR	6.0/4.0
YR	4/5	.234	rY	5.0/6.5	rY	5.0/7.0	RY	5.0/6.5
B	5/6	.211	B	5.0/7.5	gB	5.5/7.0	B	5.5/6.5
G	5/7	.183	G	6.0/9.0	G	7.0/9.0	G	6.0/9.0
R	3/7	.179	R	4.5/7.5	yR	5.0/8.0	R	5.0/7.5
P	3/6	.129	RB	4.0/5.0	RB	5.0/6.0	bR	5.0/7.0
G	3/4	.082	G	1.0/2.0	G	3.5/4.0	G	3.5/4.0
B	3/5	.052	B	3.0/5.0	B	4.0/6.5	B	4.0/5.5
A	0/0	.030	A	0.0/0.0	A	0.0/0.0	A	0.0/0.0

hue on black background as compared with white shows the tendency of dark grounds to induce illuminant hue on objects even in non-homogeneous illuminations. By its preponderant influence on adaptation level the background can determine whether objects shall be tinged with illuminant hue, the contrasting hue, or shall lose in saturation to the point of achromaticity.

Classically it was thought that non-selective backgrounds exerted only a bright-

ness contrast effect on samples but the reports in Tables II and III prove that backgrounds differing only in lightness can change not only the lightness of samples but their hue and saturation as well. This holds both for homogeneous and, to a lesser extent, for non-homogeneous illuminants. Hue and saturation depend as much upon background as lightness and can no longer be regarded as intrinsic properties associated solely with spectral reflectance of the object. Let us consider each of the three fundamental variables in relation to background when illumination is held constant.

The lightness contrast effects of different backgrounds do not appear striking unless we compare identical samples on different backgrounds *simultaneously*. Then the higher lightness of samples on darker grounds is unmistakable. When several samples are viewed on dark and light grounds *successively*, the *Os* report changes in hue and saturation rather than in lightness due to shift in lightness adaptation.

TABLE IV

AVERAGE LIGHTNESS AND SATURATION OF 18 SAMPLES IN 4 ILLUMINANTS ON 3 BACKGROUNDS WITH 4 INTENSITIES

(Each value is the average of 90 observations)

Lightness

Illum.	White bkgd.				Gray bkgd.				Black bkgd.			
	I	II	III	IV	I	II	III	IV	I	II	III	IV
R	4.33	4.86	4.19	4.18	4.99	5.03	4.30	3.92	4.77	4.47	4.61	4.22
Y	5.08	5.08	4.86	4.87	5.36	5.14	4.93	4.96	5.50	5.67	5.38	5.14
G	4.75	4.92	4.40	4.55	5.25	5.14	4.65	5.01	4.94	5.03	4.66	4.62
B	4.47	4.94	4.26	4.10	4.94	5.16	4.47	4.22	5.11	4.92	4.36	4.19
Av.	4.66	4.95	4.43	4.43	5.14	5.12	4.59	4.54	5.08	5.02	4.75	4.54
Grand Av.			4.62				4.85				4.85	

Saturation

Illum.	White bkgd.				Gray bkgd.				Black bkgd.			
	I	II	III	IV	I	II	III	IV	I	II	III	IV
R	3.33	2.86	2.32	2.86	3.50	4.20	3.99	3.42	4.77	5.97	5.05	4.55
Y	4.64	4.14	3.82	4.22	4.69	5.14	4.22	4.28	4.75	5.00	4.44	4.19
G	3.14	3.61	3.83	3.55	4.08	4.58	3.44	4.32	4.58	4.11	3.65	3.35
B	2.36	3.25	2.47	2.18	3.69	4.17	3.69	2.93	4.42	4.94	4.05	3.11
Av.	3.37	3.47	.11	3.20	3.99	4.52	3.84	3.74	4.63	5.01	4.30	3.80
Grand Av.			3.29±1.01				4.02±0.50				4.43±0.70	

Lightness is so closely correlated with reflectance of sample, which is constant, and with adaptive level of eye, which affects all samples equally, that the *Os* tend to see samples relatively constant in lightness when background reflectance is raised or lowered within very wide limits. The lightness values given in Tables II and IV bear out this generalization.

The case is different with saturation when background is changed even though all samples are subjected to the same change. Saturation effects are selective and depend upon more complicated relations. First, the samples, on the average, gain in saturation on darker grounds in strongly chromatic illumination, as shown in Table IV. The difference in average saturation between white and gray grounds is 0.73 steps, that between the gray and black grounds is 0.41 steps, and between the black and white grounds it is 1.14 steps. While only the difference between the black and

white grounds is completely statistically reliable, the differences are all in one direction and in favor of the darker grounds.

In non-homogeneous illumination some samples are more saturated on white ground, others on black ground, a fact which seems to contradict our findings with homogeneous illuminants. However, we can explain this fact through the interdependence of lightness and saturation. Colors which are better saturated at lower lightness values will be more saturated on white ground while colors which are more saturated at higher lightness values will be more saturated on black ground. Observation of identical yellows and blues on white and black grounds simultaneously bears out the correctness of our explanation. Samples which would pass muster on one ground may be unacceptable as matches on another ground.

Changes in intensity of illumination probably have less direct effect on color than any of the factors which we have discussed. The averages for lightness and saturation in Table IV for four illuminants at four intensities having a range of 1 to 96 show very small differences which are consistently in favor of the two highest intensities. Greater saturations and lightnesses are found in the two highest intensities as compared with the two lowest intensities, but considering the large range of intensities the lightness and saturation changes seem small.

In summary we can say that hue, lightness, and saturation of objects depend as much upon background reflectance and composition of illuminant as upon spectral reflectance of the object itself. For in viewing any sample we see it against a particular background and in a given illumination with, perhaps, other samples present in the field of vision. The fact that the eye can move freely over the surface of the sample plane means that we are dealing with a dynamic event whenever colors are being named or matched. Adequate specification must predict the color of objects under the dynamic conditions of viewing even though this is much more difficult than to assume that dynamic factors can be left out of account. Progress along these lines has already been made and the work is still going on.[8]

REPRESENTATION OF COLOR TOLERANCE ON THE CHROMATICITY DIAGRAM

By DAVID L. MACADAM, Eastman Kodak Research Laboratories

The maximum acceptable variations of sample colors from a standard color for some definite use are called color tolerances.[1] The purpose of this paper is to show that the coördinate system adopted by the International Commission on Illumination in 1931 is entirely adequate for the representation of color tolerances,[2] and that

[8] Formulae for the color of objects viewed in any illuminant on backgrounds of different reflectances have been devised by Judd and found satisfactory in general although a complete check with our extensive data has yet to be made. See D. B. Judd. Hue, saturation, and lightness of surface colors with chromatic illumination, *J. Opt. Soc. Amer.*, 29, 1939.

[1] Communication No. 715 from the Kodak Research Laboratories.

[2] Proceedings of the 8th Session, Commission Internationale de l'Eclairage (International Commission on Illumination, ICI) Cambridge, 1931, 19; T. Smith and J. Guild, The CIE colorimetric standards and their use, *Trans. Opt. Soc.*, 33, 1931,

any revision of the standard coördinate system based on the desire to represent or predict color tolerances more conveniently than on the present standard system would be illusory and is unnecessary.

Color tolerances must be established by agreement between the manufacturers and the users of the colored materials. Usually this agreement is secured by selecting from a group of available samples those which are acceptable for the purpose under consideration. A wise manufacturer will submit for such selection samples having various degrees of all probable variations from the standard color. The consumer, realizing that the cost of production will depend on the percentage of rejections demanded by the limits he selects, will accept all colors whose departures from the standard are not seriously objectionable. The limiting acceptable departures from the standard color will therefore depend on two important factors: (1) the ease with which the manufacturer can control the several possible kinds of variations from the standard color; and (2) the relative objectionabilities of these different kinds of variations in the opinion of the user of the colors. The first of these factors does not depend in any way on the degree of noticeability of the variations from the standard color, and the second factor depends only partially on noticeability. For some purposes, noticeable variations from the standard dominant wave-length will be much more objectionable than equally noticeable variations from the standard purity, and for other purposes the relative objectionabilities of these same variations from the same standard color will be quite reversed. The relative noticeabilities of dominant wave-length variations and purity variations are not determined by the purpose for which the color is used, and depend only on the dominant wave-length and purity of the standard color, and on the conditions of observation, including the shape and the size of the colored areas being compared, the distribution and color of the light in the surrounding field of vision, and the state of brightness and color adaptation of *O*'s eye at the time of the comparison.

Typical examples of color tolerance specifications. The A.A.R. Signal Section Specification 69-35 for railway signal glasses is expressed in terms of the ICI 1931 standard coördinate system, as shown in Fig. 1. Only the lower limit of the luminous transmittance for the 2360°K. Planckian radiator is specified for each glass, since the glasses are to be used in a manner which does not require any upper limit of transmittance, provided that the chromaticity of each glass is within the limits shown in the figure. An additional specification is required to differentiate the blue and purple signals. In practice the principal difference between these two signal colors is not one of chromaticity, but is due to the fact that the purple glass gives a dichromatic signal, *i.e.* red surrounded by blue, resulting from the chromatic aberration of the eye, whereas the blue glass, transmitting but little red light, gives only the blue signal color. This difference in the relative amounts of red light transmitted by the blue and purple glasses forms an essential part of the specification for these two signal colors. For this particular purpose, red light is defined to include the wave-length from 650 to 770 mμ, and the proportion of the red to total luminous transmission is expressed by the ratio T_r/T_w. There is no evident relation between the relative magnitudes of the tolerances in different directions on the chromaticity diagram, and the relative magnitudes of the just noticeable dif-

73-130; D. B. Judd, The 1931 ICI standard observer and coördinate system for colorimetry, *J. Opt. Soc. Amer.*, 23, 1933, 359-374; Handbook of Colorimetry, *Technology Press, Mass. Inst. Tech.*, 1936.

ferences of chromaticity in the corresponding directions.[3] These tolerances on the chromaticity and the lower limits of transmittance for railway signal glasses were determined by agreement, in which the primary considerations were availability of the glasses described, and their usefulness for their specific purpose. The exact locations of the tolerance boundaries were chosen so that they could be expressed simply in analytical form, in terms of the ICI coördinate system, and the fact that

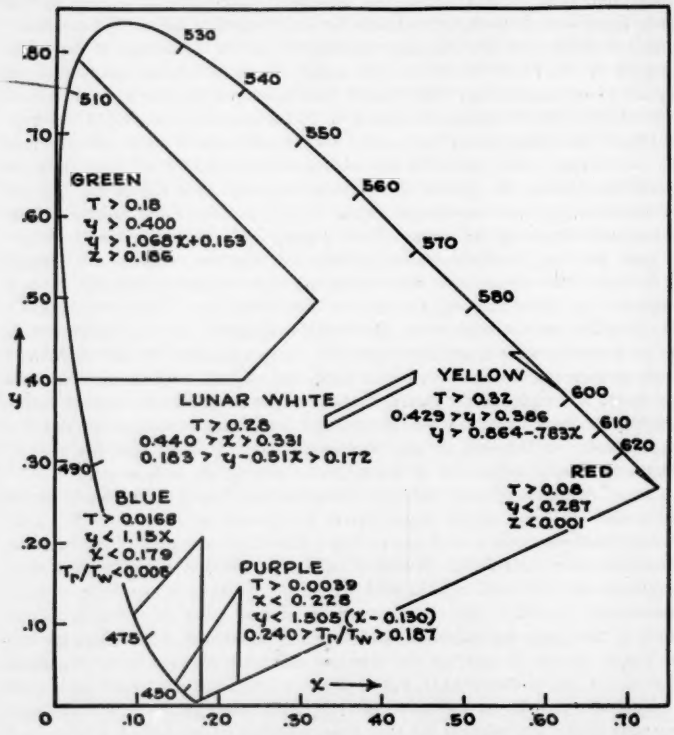
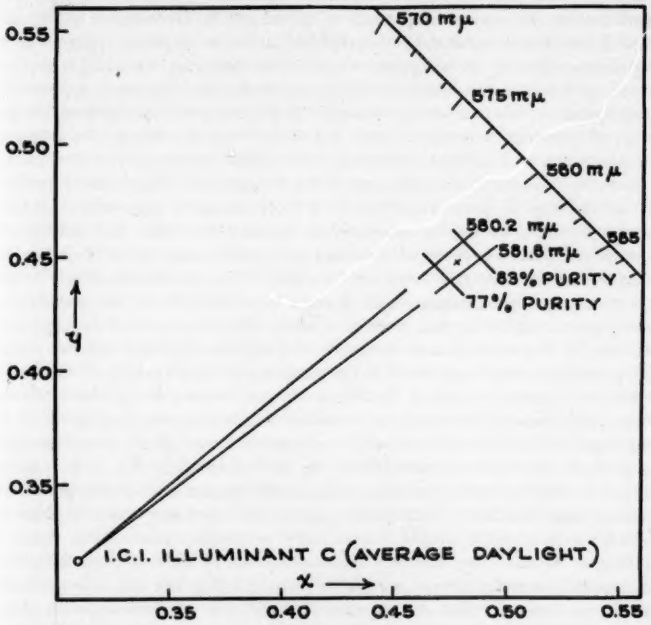


FIG. 1. CHROMATICITY DIAGRAM BASED ON 1931 ICI STANDARD COÖRDINATE SYSTEM FOR COLORIMETRY

these boundaries consist of straight lines demonstrates that convenience of specification was considered of greater importance than noticeability of color differences. Such a series of tolerance specifications could be made in any trichromatic coordinate system based on valid color mixture data. A coördinate system based also

[3] D. B. Judd, A Maxwell triangle yielding uniform chromaticity scales, *J. Opt. Soc. Amer.*, 25, 1935, 24-35; Estimation of chromaticity differences on the standard ICI colorimetric coördinate system, *ibid.*, 26, 1936, 421-426.

on just noticeable color difference data would have been of little assistance in the establishment of suitable tolerances for railway signal glasses.[4] Noticeability of color difference between signal glasses of the same kind probably would be over-



	STANDARD	TOLERANCES
DOMINANT WAVELENGTH	581 mμ	580.2 - 581.8 mμ
EXCITATION PURITY	80%	77% - 83%
DAYLIGHT REFLECTANCE	54%	52% - 56%

FIG. 2. TOLERANCE LIMITS ADOPTED FOR THE INSPECTION OF A YELLOW REFLECTING MATERIAL

emphasized in importance, perhaps at the expense of convenience, by such use of a uniform chromaticity scale coördinate system.

Fig. 2 shows the tolerance limits adopted for the inspection of a yellow reflecting material. Both upper and lower limits of daylight reflectance are included in this tolerance specification, because daylight reflectance is an important characteristic

[4] T. Smith, The colour triangle and colour discrimination, *Report of Joint Discussion on Vision, University Press, Cambridge,* 1932, 212-226; F. C. Breckenridge and W. R. Schaub, Rectangular uniform-chromaticity-scale coördinates, *J. Opt. Soc. Amer.,* 27, 1937, 226; R. H. Sinden, A further search for the ideal color system, *ibid.,* 27, 1937, 124-131; 28, 1938, 339-347.

of the colors of reflecting surfaces. In this specification, the tolerance boundaries coincide with loci of constant dominant wave-length and excitation purity compared to the ICI illuminant C. This type of limit specification was adopted because the magnitude of the just acceptable variation of dominant wave-length appeared in repeated tests to be almost independent of all acceptable variations of purity and daylight reflectance; because the just acceptable variations of purity appeared to be almost independent of all acceptable variations of dominant wave-length and reflectance; because the just acceptable variations of daylight reflectance appeared to be independent of all acceptable variations of chromaticity; and because the expression of tolerances in terms of upper and lower limits of dominant wave-length, excitation purity, and daylight reflectance corresponded quite closely to the manner in which the inspectors customarily grade the samples, and classify the rejections. That noticeability of color variations is of only secondary importance in these tolerances is confirmed by the agreement of the inspectors that they can see the maximum acceptable variations of dominant wave-length more easily than they can see the maximum acceptable variations of purity. This corresponds closely to the fact that dominant wave-length, which depends quite critically on the concentration of the pigment applied to the surface, is more difficult to control in large-scale production of this material than are purity and reflectance, which depend, respectively, primarily on the composition of the pigment and on the quality of the surface to which the pigment is applied. In this, as in most commercial problems of color tolerance, the shape of the tolerance boundary on the chromaticity diagram is determined primarily by the relative costs of production control of the several probable variations of color. Colored materials of the kind specified in Fig. 2 and having high purity and low reflectance look soiled. Such samples having low purity and high reflectance look faded. Consequently, purity and reflectance variations from the standard are more objectionable than equally noticeable variations in dominant wave-length. In this case, therefore, production control difficulties combine with objectionability to make the just acceptable variations of purity and reflectance considerably less noticeable than the just acceptable variations of dominant wave-length.

In general, it must be anticipated that, for some purposes, the size and shape of the tolerance boundary on the chromaticity diagram may depend on the luminous reflectance (or transmittance). The tolerances of photographic and photomechanical color reproduction appear to require some such dependence, which is not primarily determined by the variation of noticeability of color variations among samples having different brightnesses when the eye of the critic is maintained at a fixed state of adaptation.

Relation of noticeability to color tolerance. The examples just given are indications that noticeability of color variations is not the sole factor determining color tolerances, nor is it among the most important factors in determining the shape of the tolerance boundary on the chromaticity diagram. Only in cases where a 'perfect' color match is required, regardless of all other considerations, will noticeability be the most important consideration determining color tolerances. A perfect color match can only mean that the sample is indistinguishable from the standard under the conditions of examination of the samples. It is commonly assumed that the tolerances in such cases are proportional to the just noticeable color differences from the standard color. This assumption is reasonable provided that the just noticeable dif-

ferences have been determined with the same conditions of observation as those prevailing at the time of comparison of the standard and sample materials. These conditions include the size and brightness of the colored areas being compared, and the distribution and color of the light in the surrounding field of view both before and during the comparison. No one set of data on just noticeable color differences will therefore be satisfactory for the derivation of color tolerances in the widely varying conditions of observation with which miscellaneous materials are examined and used. Consequently, a trichromatic coördinate system designed to represent just noticeable differences in some simple fashion would be valid for only one class of observations, and would be useful for the derivation of the tolerances for 'perfect' color matches only if the samples are to be examined under the corresponding conditions. For example, a coördinate system based on just noticeable differences for samples subtending a small angle at *O*'s eye and observed with an unilluminated surrounding field might conceivably be valid for the establishment of color tolerances for signal lights for use at night, but such a system would be quite misleading if used for the establishment of color tolerances for textiles used in surroundings well illuminated with various phases of daylight. The standardization of a chromaticity diagram which conveniently represents just noticeable differences with any one set of conditions of observation would encourage the use of that data for all problems of color tolerance regardless of the applicability of the data to the conditions of observation imposed by each specific problem.

Although no one uniform chromaticity scale diagram is of sufficiently general applicability to justify its standardization in place of the 1931 ICI standard coordinate system for general use in colorimetry, each uniform chromaticity scale diagram can be used advantageously for many problems in which the corresponding conditions of observation are employed. Each uniform chromaticity scale diagram is therefore desirable as an auxiliary to the 1931 ICI standard coördinate system. The derivation of such uniform chromaticity diagrams for all important sets of conditions of observation is a task worthy of serious investigation, and encouragement from standardizing committees. The transformation of data from the standard coördinate system to each of the uniform chromaticity scale diagrams should be put in the simplest possible form for computation.[5] If such transformation methods are made available, all data expressed in terms of the standard coördinate system can be easily transformed to the most suitable uniform chromaticity scale diagram. The ambiguity and inconvenience which might result from the wide-spread use of several coördinate system in colorimetry can be avoided if all published data are expressed in terms of the 1931 ICI coördinate system as well as in terms of whichever uniform chromaticity scale system is appropriate to the conditions of the special problem.

Summary. Actual practice in the establishment of color tolerances indicates that visual examination of a representative group of samples, and agreement between the manufacturers and representative users of the colored materials is more satisfactory to all parties concerned than any theoretical deduction of tolerances from abstract experiments. Tolerances established by such agreement can be represented

[5] D. L. MacAdam, Projective transformations of ICI color specifications, *ibid.*, 27, 1937, 294-299.

just as clearly on the ICI chromaticity diagram as on any other chromaticity diagram. The ICI chromaticity diagram is recommended for such use because it has been standardized internationally and used extensively for many years, and because its use will not encourage any false simplifications of the color tolerance problem.

SPECIFICATION OF COLOR TOLERANCES AT THE NATIONAL BUREAU OF STANDARDS

By DEANE B. JUDD, Washington, D. C.

The various branches of the Federal Government often request the services of the National Bureau of Standards in the preparation and certification of color standards used in the inspection of manufactured materials or articles. Similarly, various national organizations have sought and obtained the coöperation of the Bureau in such work. In Table I are listed some of the branches of the government and some

TABLE I

ORGANIZATIONS RECENTLY RECEIVING ASSISTANCE FROM THE NATIONAL BUREAU OF STANDARDS, THE MATERIALS DEALT WITH, AND THE KIND OF WORKING STANDARDS USED

No.	Organization	Material	Working standards
1	Federal Specifications Executive Committee	paint	plaques of vitreous enamel on metal
2	Bureau of Navigation and Marine Inspection	glassware for marine signal lights	glass squares
3	Bureau of Air Commerce	glassware for aviation signal lights	glass squares
4	Department of the Interior	chinaware	chinaware
5	Post Office Department	decalcomania gold letters and numerals for Post Office trucks	decalcomania gold letters on cardboard plaques
6	Association of American Railroads	glassware for railway signal lights	glass squares
7	American Oil Chemists' Society	cottonseed oil	lovibond glasses
8	National Association of Sanitary Ware Manufacturers	sanitary ware	vitreous plaques
9	National Association of Hosiery Manufacturers	raw silk	plaques of vitreous enamel on metal
10	National Retail Dry Goods Association	kitchen and bathroom accessories	plaques of vitreous enamel on metal

national associations recently receiving such assistance, together with a description of the several materials dealt with and the kinds of working standards used.

There are usually three parts (A, B, and C) in a color specification for the purchase of materials. Part A is the specification in fundamental terms of the color or range of colors acceptable. This part of the specification includes a statement of the angular and spectral distribution of the illuminant, the angle of view, the spectral characteristics of the observer, and the coördinate system within which the color or range of color is specified. Part B is the specification of the working standards together with the tolerances which they must meet. Part C consists of the directions for

the use of the working standards. These three parts are not always given explicitly, but they are implied in every color inspection accomplished by direct visual comparison with working standards. This analysis was first carried out by Gibson and Walker in their work on standardization of railway signal glasses.[1]

The purpose of a color specification is to assure that the colors of purchased materials shall fall within a prescribed range. Sometimes the selection of such a range is governed almost wholly by possible degree of manufacturing control, but a more usual consideration is the perceptibility of the color difference between the standard and the delivered unit. The methods of stating and administering color tolerances vary with the material to be purchased and its intended purpose. The present paper will discuss these methods and will show their dependence upon psychophysical data and techniques.

(1) *The 1931 ICI system.* The 1931 ICI standard observer and coördinate system is particularly applicable to the part of the color specification in which consideration of perceptibility of the color difference is not important, that is, to Part A. It has been used in Part A of Specifications 1, 2, 3, 6, 9, and 10 listed in Table I. The color range is indicated by specifying the permissible area on the mixture diagram (usually the (x,y)-diagram) combined with stated tolerances restricting luminous transmission of samples viewed by transmitted light, or restricting luminous apparent reflectance of samples viewed by reflected light.[2]

In Fig. 1 is shown an example (Specification 6, Table I) of the use of the (x,y)-diagram to indicate permissible chromaticity ranges. This diagram illustrates Part A of A.A.R. Signal Section Specification 69-38, Railroad Signal Glasses (exclusive of Kerosene Hand Lantern Globes). The permissible areas for red, yellow, green, lunar white, blue, and purple signal colors are indicated. The equations of the straight lines forming the boundaries of the areas are also given. Thus a green pressware glass would conform to Part A of this specification if, when combined with its intended illuminant, its trichromatic coefficients (x, y, z) be such that: (1) y be greater than 0.400; (2) z be greater than 0.186; (3) y exceed $1.068x$ by more than 0.153; and (4), its luminous transmission exceed a prescribed minimum value.

It will be noted that perceptibility of one green signal as different from another is of little concern in fixing these boundaries. It is only necessary to enforce the liberal tolerance that the signal be not too yellowish to be confusable with the yellow signal nor too pale to be confusable with the lunar white, and so on. These limits were set by the signal engineers by experiments with the whole group of signals under service conditions.

Part A of Specification 7 (Table I) is given in terms of the O.S.A. observer and coördinate system,[3] which was superseded by the ICI system in 1931 wherever feasible in the work of the National Bureau of Standards. Furthermore, Part A of Specification 5, given in the uniform-chromaticity-scale (UCS) system,[4] could have

[1] K. S. Gibson and G. K. Walker, Standardization of railway signal glasses, *Signal Section Proceedings, A.R.A.* 30, 1933 (No. 5), 384.

[2] For definition of apparent reflectance consult H. J. McNicholas, Absolute methods in reflectometry, *Bur. Stand. J. Res.*, 1, 1928, 29-73.

[3] L. T. Troland, Report of committee on colorimetry for 1920-21, *J. Opt. Soc. Amer.* 6, 1922, 527-596.

[4] D. B. Judd, A Maxwell triangle yielding uniform chromaticity scales, *J. Opt. Soc. Amer.*, 25, 1935, 24-35; also *Bur. Stand. J. Res.*, 14, 1935, (no. 756), 41-52.

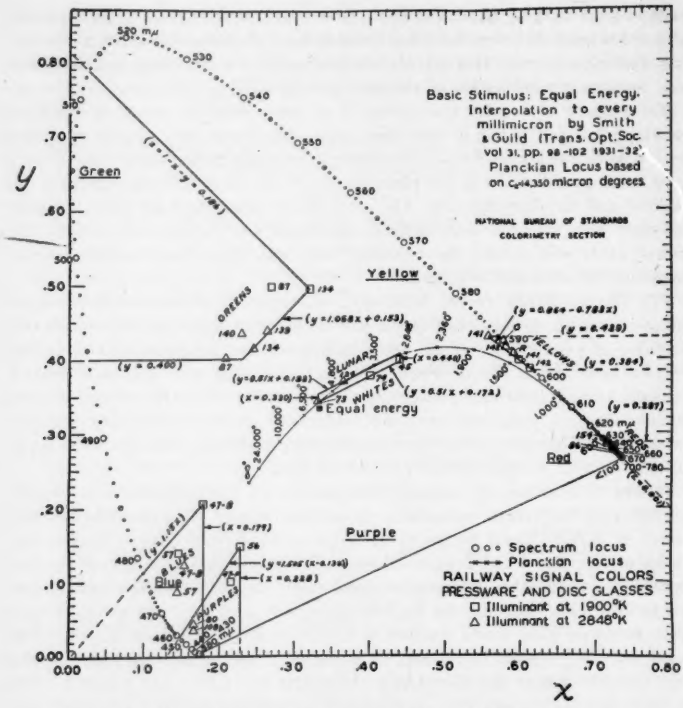


FIG. 1. MIXTURE DIAGRAM ACCORDING TO THE 1931 ICI STANDARD COÖRDINATE SYSTEM

This diagram illustrates Part A of the Association of American Railroads Signal Section, Specification 69-38, for railroad signal glasses exclusive of kerosene hand lantern globes. It is taken from their published specification and is an example of the use of the ICI system for specification of color tolerances.

been given in the ICI system, in agreement with the statement by Dr. MacAdam of its general applicability,[5] though somewhat less conveniently since Part B is particularly suited to the UCS system.

The ICI system is also used in Part B of Specifications 2, 3, 6, and 9 (Table I), but is in no case used in Part C.[6] There is no reason why each item of delivered

[5] D. L. MacAdam, Representation of color tolerances on the chromaticity diagram, this JOURNAL, *supra*, 412-418.

[6] The actual application of working standards to the color inspection of manufactured articles or materials described in Part C of a color specification is carried out at the Bureau so infrequently that the methods scarcely come within the scope of this paper. For the sake of completeness, however, it may be mentioned that

goods should not be compared with the working standard by means of the ICI observer instead of by direct visual inspection except that the required spectrophotometry plus computation is usually much the more time-consuming. Furthermore, existing spectrophotometers are not well adapted to the measurement of many manufactured articles; for this reason Part A of Specifications 4, 8, and 9 is not expressed in any fundamental system. These specifications thus have to depend, for the present, upon the integrity of material master standards of color.

(2) *Standard and tolerance sample.* Whenever the purpose of a color specification is to make sure that delivered goods depart from a standard color by as little as possible, Part A will be the specification of a single standard color, and Part B will include a color tolerance for the working standards. If, in addition a master standard in material form is available, the color tolerance for the working standards frequently takes the form of a tolerance sample. To apply such a tolerance the observer has to judge whether the color difference between the purported working standard and the master standard is greater or less than that between the tolerance sample and the master standard. In Dr. Boring's terminology, this is a sense-distance judgment corresponding in the critical case to the method of bisection (AB = BC.[7] In Dr. Newhall's terminology it is the "ratio method" in which the critical value of the ratio is unity.[8]

The directness of the method and its aptness for the central purpose of the color specification commend it; its only disadvantage is the lack of precision and reproducibility. It is common for even a trained observer to change his judgment by a factor of 2; and one trained observer of the Bureau staff will differ from another frequently by factors of 3 or 4.

This method was used until recently in Part B of Specification 8 (Table I), and it was also used in Part B of Specification 10, except for critical samples, because of its relative simplicity and rapidity.

(3) *Unit of surface-color difference.* In an effort to develop a method of setting color tolerances which would combine the aptness of the sense-difference judgment

---

when Part A specifies a single color, Part C usually states the observing conditions (field size, illuminant, and so on) for the comparison of the working standard with the manufactured goods, and says that any appreciable difference warrants rejection (*e.g.* Specifications 8 and 10). If Part A specifies a color range that is essentially a one-dimensional continuum, Part B will specify working standards representing the limits of this range, and Part C will state that the color of the manufactured goods shall fall between the limits exemplified by the two working standards. The working standards of Specification 7 and some of Specification 3 are used in this way. If, however, the range of color specified by Part A is not a simple one-dimensional continuum, Part B will often specify more than two limit standards, and Part C will usually require the inspector to make a more complicated judgment such as "not greener than the green-limit standard" and "not lighter than the light-limit standard" and so on. The difficulty of these judgments arises from the necessity of the inspector having to respond to the algebraic sign of a difference in one attribute of color in the presence of fairly large differences in other attributes; that is, he has, for example, to judge the direction of a hue difference in the presence of a fairly large saturation or lightness difference, or both. These more complicated responses are required by Part C of Specifications 2, 3, 5, and 6.

[7] E. G. Boring, The psychophysics of color tolerance, this JOURNAL, *supra*, 384-394.

[8] S. M. Newhall, The ratio method in the review of the Munsell colors, this JOURNAL, *supra*, 394-405.

with the precision and reproducibility afforded by the ICI coördinate system, a program of research has been prosecuted at the National Bureau of Standards for several years. The research is not completed, but the preliminary techniques are sufficiently successful that the unit of mensuration defined by them has been called the NBS unit of surface-color difference. The techniques are furthermore applicable to many samples which cannot be measured by existing spectrophotometers; this applicability results from the design of reflectometers and chromaticity-difference colorimeters, both visual and photoelectric, suited to cylindrical as well as flat samples, and to a considerable range of sample size.[9]

(a) The uniform-chromaticity-scale triangle. The first step in this research program leading to the general derivation of scales of uniformly perceptible chromaticity intervals was taken by Priest and Brickwedde in 1926 when they determined sensibility to chromaticity differences among stimuli closely resembling sunlight.[10] Before a complete account of these determinations was published in 1938,[11] they had already been corroborated by three independent investigations.[12] The method used in all of these determinations was direct setting of the just noticeable difference (jnd).

The second step was a critical review of the literature of chromaticity sensibility which resulted in the discovery of an empiric relation with the O.S.A. colorimetric coördinate system.[13] The literature studied consists of about 80 papers, of which more than 20 supplied useful quantitative data, some based upon direct setting of the jnd, others on the precision of setting for a chromaticity match.

Other paths of the mixture diagram were then explored by Judd[14] and Tyndall,[15] and in 1935 these new data were combined with the old as guides in the derivation of a projective transformation of the ICI system whose mixture diagram

---

[9] R. S. Hunter, A null method photoelectric reflectometer, *J. Opt. Soc. Amer.*, 26, 1936, 225-230, also *Bull. Amer. Ceramic Soc.*, 15, 1936, 79-80; Problems in the development of a multipurpose reflectometer, *J. Opt. Soc. Amer.*, 27, 1937, 225-235, also *Bull. Amer. Ceramic Soc.*, 16, 1937, 90-94; D. B. Jull, A subtractive colorimeter for the measurement of small chromaticity differences between surfaces of moderate spectral selectivity of reflectance, *J. Opt. Soc. Amer.*, 26, 1936, 225-230; Specification of uniform color tolerances for textiles, *Textile Res.*, 9, 1939, 253-263, 292-308.

[10] I. G. Priest and F. G. Brickwedde, The minimum perceptible colorimetric purity as a function of dominant wave length with sunlight as neutral standard, *J. Opt. Soc. Amer.*, 13, 1926, 306-307.

[11] Priest and Brickwedde, Minimum perceptible colorimetric purity as a function of dominant wave length, *J. Opt. Soc. Amer.*, 28, 1938, 133-139, also *Bur. Stand. J. Res.*, 20, 1938, (no. 1099), 673-682.

[12] D. M. Purdy, On the saturations and chromatic thresholds of the spectral colours, *Brit. J. Psychol.*, 21, 1931, 283-313; L. C. Martin, F. L. Warburton, and W. J. Morgan, Determination of the sensitiveness of the eye to differences in the saturation of colors, *Spec. Rep. Ser. Med. Res. Coun. Lond.*, 1933, (no. 188), 1-42; W. D. Wright and F. H. G. Pitt, The saturation-discrimination of two trichromats, *Proc. Phys. Soc. London*, 49, 1937, 329-331.

[13] Judd, Chromaticity sensibility to stimulus-differences, *J. Opt. Soc. Amer.*, 22, 1932, 72-108.

[14] D. B. Judd, Sensibility to color-temperature change as a function of temperature, *ibid.*, 23, 1933, 3-14; Saturation scale for yellow colors, *ibid.*, 23, 1933, 35-45.

[15] E. P. T. Tyndall, Chromaticity sensibility to wave-length difference as a function of purity, *ibid.*, 23, 1933, 15-24.

plotted in triangular coördinates was intended to indicate by distance the perceptibility of differences between color stimuli. This is the uniform-chromaticity scale (UCS) system; it is the basis of the present NBS unit of surface-color difference. Fig. 2 shows the mixture diagram of the UCS system. Parts A and B of Specification 5, Table I, which is purely a chromaticity specification, make use of the UCS system.

(b) Derivation of the formula for color difference. The basic idea by means of

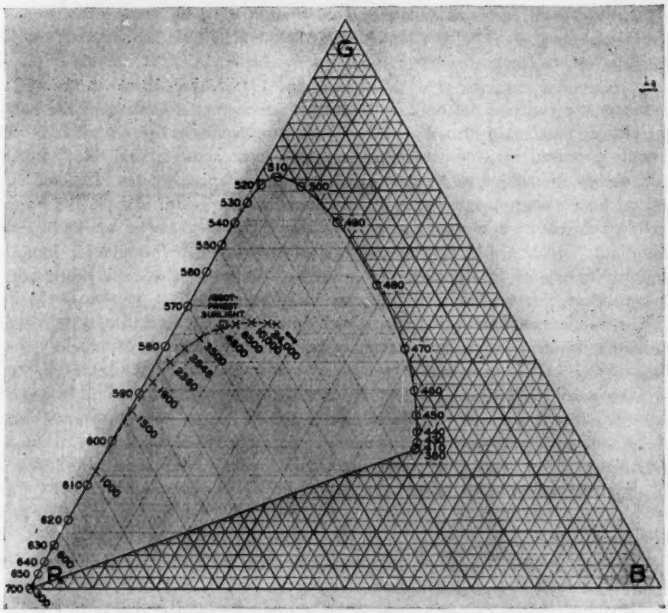


FIG. 2. THE UNIFORM-CHROMATICITY-SCALE TRIANGLE
Wave-length along the spectrum locus is indicated in mμ; color temperature along the Planckian locus is given in degrees Kelvin.

which this formula for size of differences among surface colors was derived is that Euclidean geometry can be applied; that is, it is assumed that it is possible to represent surface colors by points so arranged as to make up a solid figure, known as the surface-color solid, within which Euclidean distances between points are proportional to the size or perceptibility, ΔE, of the color difference between the colors represented by the points.

The first assumption on which the formula depends is that surface colors of the same lightness (Munsell term is value) can be represented on horizontal planes in the surface-color solid, and that variations in lightness at constant chromaticity

are to be represented along lines perpendicular to the constant-lightness planes.[16] If we have, therefore, two colors, E_1 and E_2, having lightnesses, L_1 and L_2, and chromaticities, C_1 and C_2, this assumption yields by ordinary geometry the familiar distance formula: $(E_1-E_2)^2=(L_1-L_2)^2+(C_1-C_2)^2$ or, in the difference notation which we shall use hereafter:

$$\Delta E^2 = \Delta L^2 + \Delta C^2 \quad [1]$$

The second assumption is that the true relation connecting lightness of the surface color with the luminous apparent reflectance, A, of the surface departs only by a negligible amount from the square-root relation; that is, we assume that:

$$L = k_1 A^{1/2} \quad [2]$$

or, in difference form

$$\Delta L = k_1 \Delta (A^{1/2}) \quad [2a]$$

Lightness is a function not only of the luminous apparent reflectance of the sample but also a function of that of the background against which it is viewed.[17] The formula proposed by Godlove is probably the most accurate available;[18] but we have chosen the simple relation [2] as a convenient approximation for both light and medium backgrounds because it is desired for the present to have the formula for color difference apply not to any particular background but to average viewing conditions. Equation [2] is a special case of the Godlove formula for luminous apparent reflectance of the background equal to unity; it also falls nearly within the spread of observers for a middle gray background.[19]

The third assumption deals with the use of the uniform-chromaticity-scale triangle. Although the experimental data from which this particular transformation of the standard 1931 ICI colorimetric coördinate system was derived are chiefly aperture colors of small angular extent (confined to the macula lutea) it is assumed that this triangle applies sufficiently closely to differences between surface colors of the same lightness viewed so as to subtend a large visual angle (not confined to the macula). This assumption would permit us to write for chromaticity difference, ΔC, between surface colors of any one lightness:

$$\Delta C = k_3 \Delta S \quad [3]$$

where:

$\Delta S^2 = (r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2$, the trilinear coördinates of the first color being (r_1, g_1, b_1) and those of the second (r_2, g_2, b_2).

The fourth assumption has to do with the experimental fact that, under ordinary conditions of inspecting samples for color, a given chromaticity difference corresponds to a larger trilinear-coördinate difference for dark samples than for light. It is assumed that equation [3] extended to colors of all lightnesses should be written:

$$\Delta C = k_3 A^{1/4} \Delta S \quad [4]$$

---

[16] Chromaticity of a chromatic color is determined by the hue and saturation attributes of the color; see L. A. Jones, Colorimetry: Preliminary draft of a report on nomenclature and definitions, *J. Opt. Soc. Amer.*, 27, 1937, 207-213.

[17] E. Q. Adams and P. W. Cobb, The effect on foveal vision of bright (and dark) surroundings, *J. Exper. Psychol.*, 5, 1922, 39-45; E. Q. Adams, A comparison of the Fechner and Munsell scales of luminous sensation value, *J. Opt. Soc. Amer.*, 6, 1922, 932-939.

[18] I. H. Godlove, Neutral value scales: II. A comparison of results and equations describing value scales, *ibid.*, 23, 1933, 419-425.

[19] A. E. O. Munsell, L. L. Sloan, and I. H. Godlove, Neutral value scales: I. Munsell neutral value scale, *ibid.*, 23, 1933, 394-411.

This formula would state that a deep 'red' is not a perceptibly different color from a deep 'green,' provided the surfaces exemplifying them have sufficiently low luminous apparent reflectances. The reason why the eye could not detect the color difference between two such samples (both of which would appear black) is that the relatively much stronger stimulation of neighboring retinal areas prevents it by irradiation and contrast effects. Equation [4] like equation [2] holds, therefore, most closely for inspection of samples under the ordinary condition that the surrounding field is lighter than dark samples, but if precautions were taken to have no part of the retina receive greater illumination than the part on which the image of the two samples falls, then equation [4] would not hold; on the contrary, equation [3] would probably apply to colors of all lightnesses without restriction. The reason for choosing the fourth root of the luminous apparent reflectance, A, in equation [4] rather than some other function of A is first because of its simplicity, and second because the samples of the Munsell system approximately follow this relation.

By substitution of [2a]' and [4] in [1], for $k_1 = 120$ and $k_2 = 600$, we obtain the formula for size of color difference which has been tested in a preliminary way for conformity to the results of direct visual estimates.

$$\Delta E = 600\,[A^{1/2}\Delta S^2 + 0.04\Delta(A^{1/2})^2]^{1/2} \quad \ldots\ldots\ldots\ldots\ldots\ldots [5]$$

The ratio of the constants k_1 and k_2 has to be adjusted so as to evaluate lightness differences relative to chromaticity differences in the same way as the average inspector weights them. The ratio 120:600 is a tentative choice based on a few preliminary experiments for samples (such as textiles and papers) which can be compared with a dividing area either non-existent or very narrow. It is known that the presence of a conspicuous dividing area interferes with the perception of chromaticity differences,[20] but preliminary trials indicate that the perception of lightness differences is even more reduced. For a narrow dividing area contrasting with the samples by about the lightness difference betwen black and mid-gray (such as is common in the comparison of two plaques of vitreous enamel on metal, Specifications 1, 8, and 10), better agreement with direct visual estimate is found by taking $k_2 = 90$ or 100 instead of 120; separation by a conspicuous broad design of the areas to be compared requires k_2 as small as 30 or 40 for good agreement.

The sum of the constants, k_1 and k_2, determines the size of the unit of color difference; it has been adjusted with the intention of choosing a unit color difference of such size that smaller differences can be neglected in the average commercial transaction involving a color match but larger differences cannot. This adjustment refers to observation of fairly large samples (say 10 cm. squares at a distance of 25 cm.) and is also tentative; it gives 120 units of color difference from black to white, and about 300 units between the strongest possible red and the strongest possible green for samples of daylight apparent reflectances of 0.10.[21] This makes the surface-color solid a rather squat figure, the maximum horizontal dimension being nearly 3 times that of the maximum vertical dimension.

Formula [5] must be restricted to matt surface colors. With glossy samples,

---

[20] Judd, Precision of color-temperature measurements under various observing conditions: A new color comparator for incandescent lamps, *Bur. Stand. J. Res.*, 5, 1930, 1161-1177.

[21] D. L. MacAdam, Maximum visual efficiency of colored materials, *J. Opt. Soc. Amer.*, 25, 1935, 361-367.

under the usual inspecting conditions, most specularly reflected light is directed away from the eye of the inspector, but with this customary precaution there still remains enough specularly reflected light to interfere significantly with the perception of color differences between dark samples. There is less interference for samples of medium lightness and little or none for near-white samples. Improved agreement for glossy samples has been obtained by expanding the constant, 600, into the following function of the specular gloss, G_s, and luminous apparent reflectance, A, of the samples: $(600\ AB_o)/(AB_o + 0.05\ G_sB)$ where B_o is the brightness which a uniform hemisphere over the sample would need have to give the sample its actual illumination, and B is the brightness of the area imaged in the sample.[23] It will be noted that if the area imaged in the sample be completely dark $(B = 0)$ or if the sample be matt $(G_s = 0)$, this function reduces to the simple constant, 600. For inspection in an ordinary room near a window by light from the sky, values of $B = B_o/2$ are common; so for typical inspection of glossy vitreous plaques $(G_s = 1)$ the function becomes: $600A/(A + 0.025)$. This function departs but slightly from 600 for near-white samples $(A > 0.70)$ but drops to 300 for dark samples for which $A = 0.025$. This accords approximately with our preliminary results.

(c) Use and test of the NBS unit. The NBS unit of surface-color difference has been used in Part B of Specifications 4, 8, and 10. Several hundred differences between these vitreous enamel samples have been evaluated in this way and have been checked by comparison with direct visual estimates of the sense differences (the ratio method). This study is continuing. Some consistent discrepancies by about a factor of 2 were discovered in this way which suggested that the third assumption, the one dealing with applicability of the UCS system to surface colors of large angular extent, was not wholly justified. The amount of the discrepancy is about equal to the uncertainty of judgments obtained by a single observer by the ratio method, and was considered encouragingly small.

However, a coöperative study of the formula with the Du Pont Technical Laboratory by means of samples dyed for that purpose, corroborated these discrepancies, and showed for certain deep reds and blues that the formula was in error through failure of the third assumption by about a factor of 4. These experiments are described in a recent article including details of construction and calibration of the colorimeter used,[25] and also a comparison of this formula with the Nickerson Index of Fading now being studied by a subcommittee of the American Society for Testing Materials.[24]

An analysis of the nature of the discrepancies shown indicates that a considerably improved agreement could be obtained by moving the daylight point (near 6,500°K, see Fig. 2) away from the spectrum at 570 mμ to a point about twice as distant as the present point, and at the same time shortening the distance of the daylight

---

[23] Judd, Uniform tolerances for surface-color specification, *ibid.*, 28, 1938, 52-56.

[25] Judd, Specification of uniform color tolerances for textiles, *Textile Res.*, 9, 1939, 253-263, 292-308.

[24] Dorothy Nickerson, The specification of color tolerances, *Textile Res.*, 6, 1936, 505-514; How can results of fading tests be expressed, *A.S.T.M. Stand. Tex. Mat.*, 1936, 238-241.

point from the long-wave extreme of the spectrum. Both of these changes could be made by an homogeneous, linear transformation; so formula [5] would not be changed in form, but the coördinate system within which the term, ΔS, is evaluated would be different. This suggested new coördinate system would not apply so well to chromaticity-sensibility data on aperture colors of small angular extent (confined to the macula lutea) as the present uniform-chromaticity-scale system; but it is probably significant that both the increased distance between the daylight point and the spectrum at 570 mμ and the decreased distance to the long-wave extreme of the spectrum are consistent with the use of a large retinal area not confined to the macula lutea. Note that a 10 cm. sample viewed at the distance of most distinct vision (about 25 cm.) would subtend an angle of 22°; this is more than 5 times the diameter of the average macular pigment.

Whether the suggested new coördinate system will result in the expected improved agreement for deep reds and blues and, at the same time, avoid introducing discrepancies for other colors, remains to be determined by trial. If it should fail, the more complicated and more flexible formulation by Sinden,[25] which also is consistent with an Euclidean arrangement, could be tried next; then completely unrestricted (non-Euclidean) forms.

Summary and conclusions. The various parts of a color specification to be administered by working standards have required different methods for color-tolerance specification. Choice of method is also affected by the article whose color is specified and by the instruments used. Three chief methods are discussed: (1) the standard ICI coördinate system; (2) material standard and tolerance sample; and (3) the NBS unit of surface-color difference.

The ICI system applies aptly to color specifications not requiring account of the perceptibility of differences; it yields a precise and reproducible specification of color tolerance in fundamental terms. The use of a material standard and tolerance sample requires a sense-distance judgment, and within the limitations of precision of that judgment is perfectly adapted to insure tolerances of uniform perceptibility. The NBS unit of surface-color difference is intended to combine the precision and reproducibility of a fundamental system with the aptness of the sense-distance judgment by means of a tolerance sample. The formula for color difference defining this unit makes use of sensibility data obtained over a period of years by psychophysical methods, and in effect interpolates between the various series of stimuli for which color-difference sensibility is known. The NBS unit of surface-color difference suffers from two disadvantages: first, the psychophysical data on which it is based are so small a proportion of the total required for reliable evaluation of the complete surface-color solid that attempts to use the unit still reveal ways in which the interpolations yielded by the color-difference formula may be significantly improved; and, secondly, the unit varies with observing conditions so that the color-difference formula defining it is necessarily rather complicated. The NBS unit is already successful enough to justify its use in some standardization work, but our research continues to be directed toward its improvement. It should be emphasized therefore, that changes in definition of the unit are to be expected, but it is believed that

[25] R. H. Sinden, A further search for the ideal color system: I. A new mechanicographical method, *J. Opt. Soc. Amer.*, 27, 1937, 124-131; II. A reconsideration of the Helmholtz line element, *ibid.*, 28, 1938, 339-345.

a practical unit of surface-color difference will result from our study, and that the definition of the unit will be no more complicated than the tentative form discussed here.

INDUSTRIAL COLOR TOLERANCES

By Isay A. Balinkin, University of Cincinnati

The scope of the subject of industrial color tolerances embraces such a wide field that we shall discuss only those phases, defined in the previous papers, which have direct application to the problems of evaluations of color differences, tests on consumers acceptance, and production control of color uniformity. Since the author's connections for the past three years have had to do with ceramic glazed wall tiles, problems of industrial color tolerances will be discussed in their relation to that particular industry. In so far as the principles discussed are of fundamental nature, they will be found to be also applicable to other industrial or commercial products.

The ceramic industry is a good proving ground for finding out the usefulness and applicability of recently developed methods and techniques for measuring color and color differences. There are several reasons why such is the case. (1) The variety of available colors covers a wide range of hues with large variations in brightness and saturation. (2) The colors are permanent. This is particularly important if colors are to be measured on different instruments, or the precision of the instrument itself is to be determined. (3) The surface texture of tiles is fairly uniform. (4) In the present state of manufacturing and color control there is a large color variability from tile to tile. (5) The size of the tiles being 4¼ x4¼ in. makes them very easy to handle and their color can be measured on standard instruments. (6) And finally, there seems to be a rather consistent demand for a close color matching between the tiles used in the same installation.

(1) Standard conditions. An important factor which gave impetus to the considerable amount of activity going on at present in the field of color measurements is the group of resolutions adopted by the International Commission on Illumination which met in Cambridge, England, in 1931.[1] For surface-color specification it was found necessary to establish certain standard conditions dealing with the light source, visual response of a standard observer, viewing conditions, and proper coordinate system. The adoption of standard conditions made it possible to specify uniquely any color in terms of trichromatic coefficients. At the present time these standards are adopted by scientific workers in colorimetry all over the world. Useful as such specifications are, they ought to be considered only as the first step toward solutions of many industrial problems where color is an important attribute of manufactured goods. The I.C.I. system of color specification deals primarily with color *identity*. Two colors are defined and appear to be identical under standard conditions when their trichromatic coefficients are numerically the same. Nothing

[1] D. B. Judd, The I. C. I. standard observer and coördinate system for colorimetry, *J. Opt. Soc. Amer.*, 23, 1933, 359-374. T. Smith and J. Guild, The C. I. E. colorimetric standards and their use, *Trans. Opt. Soc., London*, 33, 1931, 73-134.

is said about the evaluation of color variability or color differences. In fact, the possibility of expressing color differences as numerical quantities is considered by many as impossible or meaningless. Yet, if the definition of color is based upon a definite set of physical operations,[2] there is no reason why the definition of color differences could not be based also upon some other set of physical operations.

The I.C.I. system of color specifications is frequently referred to as the *objective* method of color specifications, whereas the question of numerical evaluation of color difference is considered to deal with purely *subjective* aspects which do not lend themselves to mathematical computations. It must be remembered, however, that the specification of color identity does involve an element of subjectivity which comes from the incorporation of psychophysical data into the I.C.I. specifications. Since our present mode of color specification involves a human element determined by the response of the retina to different radiant energy as a function of wave-length, it may be possible to establish also a scale of color differentiation determined by the response of the human eye to stimulus-differences. A quotation from a paper by Hardy stresses the fact that this problem is by no means a simple one. "If it happens that they (two colors specified in I.C.I. system) would not look alike, it is ordinarily impossible to specify the degree of their difference in terms of the color difference that a normal observer could just perceive. The determination of just perceptible differences for a large gamut of colors under a wide variety of observing conditions is a task whose magnitude dwarfs the development of the science of colorimetry. Nevertheless, it is an investigation that can profitably be undertaken, now that the imperishable language of color has been devised."[3]

We are facing now a situation in which the standard conditions for specifying color-sameness enumerated by I.C.I. will have to be supplemented by other factors which affect the appearance of color differences. Among these we may enumerate the following: (1) size of the viewing field; (2) line of demarcation between the colors; (3) intensity of illumination; (4) color of the background; (5) and finally, condition of observer's light adaptation. It is only when we give complete specifications for these additional requirements that we can even begin to talk intelligently about the problem of color differences.

In considering color differences of tiles each one of the above factors was predetermined by the actual conditions of wall tile in use. The viewing field was taken as corresponding to the angle subtended by two tiles under comparison at a distance of approximately 18 in. The line of demarcation was limited to a condition resulting when the tiles were placed in contact with each other. Observation was by daylight illumination in the neighborhood of 100–200 foot-candles. The background was neutral gray having approximately the same luminous apparent reflectance as the samples compared. The observers were light adapted.

(2) Visual estimation of color differences. The possibility of establishing surface color tolerances upon a firm scientific basis leads to two requirements: first, physical measurement of color and its evaluation as a numerical quantity; and secondly, establishment of a correlation between color differences measured physically and visual perception of such color differences. The development of physical instruments

[2] A. C. Hardy, Handbook of colorimetry, *Mass. Inst. Tech.*, 1936, 1-8.
[3] Hardy, The physical basis of color measurement, *J. Appl. Phys.*, 8, 1937, 233.

for measuring color and the adoption of the standard I.C.I. specifications gave the necessary essentials to satisfy the first requirement.

The second requirement is still in the process of development. In fact, some men of science even doubt whether such correlation can be established at all. The question is, whether color differences can be measured and, when expressed numerically, to correspond to the sense perception of such color differences. On purely philosophical or philological basis there may be a valid objection to the use of the word 'to measure' as applied to sensation, rather than to stimulus. If the true meaning of measuring is to depend upon determination of space-time coincidences, then of course there could be no meaning in speaking of measuring color differences, unless it were possible to find some process by means of which space coincidences could be established. Such an objection was raised by Guild,[4] in his criticism of the paper read by Richardson[5] during a symposium on vision sponsored by the Physical and Optical Societies in England.

From a practical standpoint such differences of opinion are only of secondary importance. The cardinal questions are: first, whether numerical magnitude of color differences can be estimated by direct visual observation; and secondly, whether there is an agreement between various observers concerning the magnitude of these color differences. The only way in which these questions can be answered is by a direct experimental test on a representative group of observers. It ought to be clear, therefore, that even before any attempt is made to discover a mathematical function applicable to the evaluation of color differences, we must be sure that both questions raised above can be answered in the affirmative.

Let us consider three surface colors designated by E_1, E_2, and E_3, which differ from each other in all three attributes: hue, lightness, and saturation. If we assign a unit value as the difference in color sensation between E_1 and E_2, that is $E_1 - E_2 =$ 1 unit of color difference, is it possible to give a numerical visual estimation of color differences between E_1 and E_3 or E_2 and E_3 in relation to the color difference between E_1 and E_2? In other words, can an observer, after viewing and comparing these three samples, state that $E_1 - E_3$ is equal to, say, 5 units and $E_2 - E_3$ is equal to, say, 3 units of color difference? And, further, if one observer would give such numerical values for his visual estimation, to what extent would they agree numerically with similar visual estimations of other observers?

In order to answer these questions the following experimental procedure was followed. Thirty-eight green tiles were selected at random and their spectrophotometric reflectance curves were made on a Hardy recording spectrophotometer. The color of these tiles varied from yellow green to blue green, from light to dark, and they were of various degrees of saturation. Trichromatic coefficients were computed from these curves for Illuminant C and each color was plotted on the I.C.I. mixture diagram. The apparent luminous reflectance of these colors was also designated by placing their values next to every point on the chromaticity plane. Every color was then designated by a consecutive number from 1 to 38, according to its increasing brightness. Five different colors were then selected to give a fairly wide range of variations in hue, brightness, and saturation, their numbers were

[4] J. Guild, Color standardisation, *Nature*, 129, 1932, 453-455.

[5] L. F. Richardson, The measurability of sensations of hue, brightness or saturation, *Rep. Joint Discussion on Vision*, 1932, 112-116.

12, 21, 24, 26, and 37. Fig. 1 is a reproduction of their spectral reflectance curves measured in the Colorimetry Section of the United States Bureau of Standards. Table I gives the trichromatic coefficients, dominant wave-lengths, luminous apparent reflectance, and purity computed by 10 mμ method for I.C.I. illuminant C.

The color difference between tile 37 and 26 appeared to be smaller than any other pair in this group of five colors and it was, therefore, assigned a value of

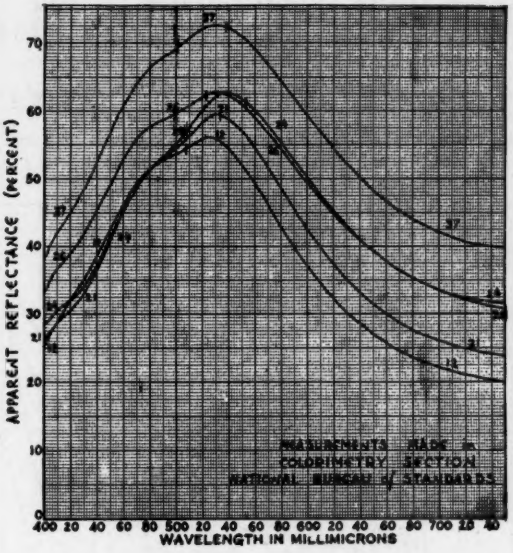


FIG. 1. SPECTROPHOTOMETRIC REFLECTION CURVES FOR FIVE GREEN COLORS USED IN ESTIMATION OF COLOR DIFFERENCES

Vertical cross-line on each curve indicates dominant wave-length for I.C.I. illuminant C.

one unit of color difference. The author proceeded then to compare color differences between all other color-pairs with this standard pair. The question proposed was: If the color difference between 37 and 26 is one, what is the color difference between 37 and 12? It was the problem of estimating directly the ratio of color difference between 37 and 12 in relation to the color difference between 37 and 26. In the form of an equation in which E stands for color and the subscript indicates the reference number of any one of the colors compared we will have $(E_{37}-E_{12})/E_{37} - E_{26} = (E_{37} - E_{12}) = R_{37\text{-}12}$ since $E_{37} - E_{26} = 1$, where $R_{37\text{-}12}$ represents the numerical ratio thus estimated.

Such a process of color difference estimation is the ratio method. With five different colors, n, the maximum possible number of pairs to be compared, N, is given by the formula: $N = n\ (n - 1)/2$, where for n equal to 5, $N = 10$. Since one

TABLE I

I.C.I. Specifications for 5 Green Colors Computed by 10mμ Method for Illuminant C

Color No.	Trichromatic coefficients x	y	z	Luminous apparent reflectance A%	Dominant wave-length (λ mμ)	Saturation P%
12	.2870	.3443	.3688	47.39	506.8	7.6
21	.2999	.3546	.3454	51.17	533.4	8.1
24	.3122	.3615	.3263	56.14	553.6	12.9
26	.2995	.3422	.3583	56.34	523.6	5.1
37	.3044	.3431	.3525	66.24	538.0	6.0

pair is already assigned a unit of color difference, only nine additional pairs have to be estimated.

Color differences estimated by the author fell between one and six times as large as the standard unit color difference. Two other *Os*, D.B.J. and R.S.H.,

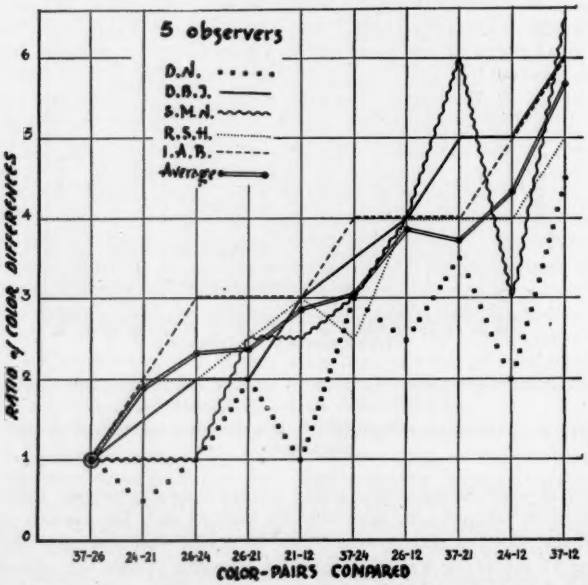


Fig. 2. Numerical Estimations of Color Differences by Ratio Method on Five Green Colors
Color-pairs are arranged according to ranking of the three first observers D.B.T., R.S.H., and I.A.B.

immediately participated in this test and it was rather surprising to find that in two instances the numerical estimates of all 3 *Os* were exactly the same, and that in 5 instances at least 2 *Os* had the same numerical ratio for color differences. The estimates of all the *Os* were made, of course, without the knowledge of the numeri-

cal values of any of the preceding *Os*. The rank of the increasing magnitude of color differences was chosen according to the estimates of these 3 *Os* which together with estimates of 2 other *Os* are shown in Fig. 2. Similar tests were made on 55

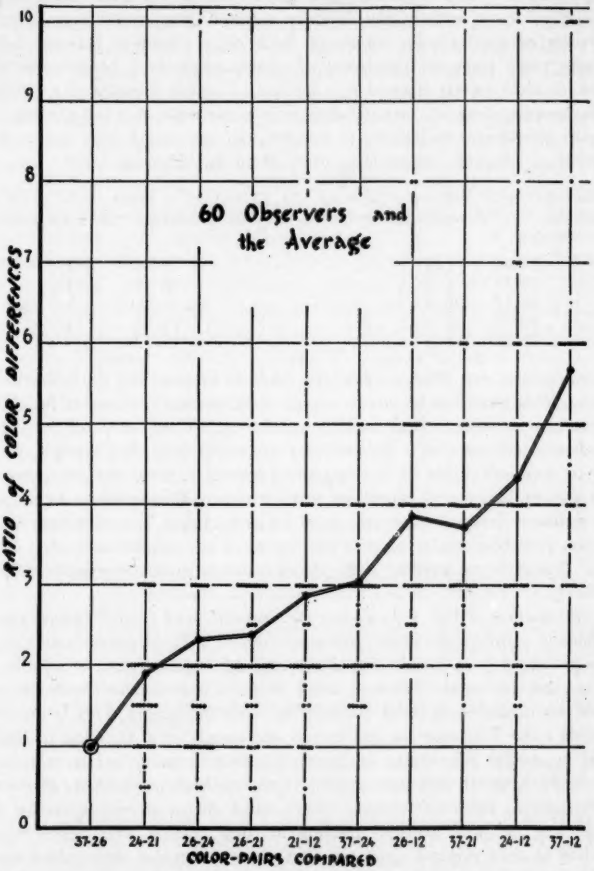


FIG. 3. NUMERICAL ESTIMATIONS OF COLOR DIFFERENCES BY RATIO METHOD ON FIVE GREEN COLORS (60 *Os*)

other *Os* and the results are presented in Fig. 3. Every point in this graph represents a single numerical estimate of color difference for a pair of colors shown along the abscissa.

In general there was no difficulty in obtaining such estimates. It was clearly

explained to every *O* that color differences were to include the three attributes of color, *i.e.* differences in hue, lightness, and saturation, all combined in a single numerical estimation in relation to a standard color difference. Some of the *O*s did experience a reluctance in naming a specific number for such ratios, although they could point out without any difficulty pairs of colors which they thought to show larger or smaller color variations. Such color variations between different pairs were made from the standpoint of acceptance as to a better color match. In some instances the *O*s resorted to a method of mental counting of a number of equally appearing intervals corresponding to a chosen unit of color difference. The time taken by different individuals to complete the test ranged from 10 to 30 min. The record of estimates was made by every *O* on the following form:

If the total color difference between tile 37 and 26 is taken to be equal to 1, what number would represent, according to your estimation, the total color difference between:

37-26	(1)	37-24	( )
37-12	( )	26-24	( )
26-12	( )	24-21	( )
37-21	( )	24-12	( )
26-21	( )	21-12	( )

After the card was filled a check was made to be sure that the estimates were consistent. This was done by cross comparison of numerical estimates for different color-pairs, to see whether their ranking was the same when compared pair by pair, rather than in reference to a standard unit color difference. For example, if 37-21 was given a ratio of 4 and 24-12 was given a ratio of 3, these two pairs were intercompared to be sure that *O* considered that the color differences were representative of his estimates between these two pairs taken by themselves. Although in some cases such cross-checking resulted in corrections of the original estimates, in every instance *O* was better satisfied with his estimates as a true representation of his opinion.

An examination of Fig. 3 shows that the distribution of color-difference estimates is sufficiently consistent to permit a computation of a fairly representative average. This is indicated by a double line drawn through solid circles. It must be said, however, that in certain instances there exists a considerable deviation among different *O*s in their numerical estimates of color differences. This is particularly true when color difference for one pair is due mainly to a variation in lightness, whereas in another pair it may be due mainly to a variation in hue or saturation. It was definitely shown that some of the *O*s had a tendency to give more consideration to differences in lightness, whereas others relied almost entirely upon the differences in hue. *O*s with previous training in color grading were very determined in holding to their original estimates, even when shown that their values were far from the averages. About half of the *O*s had some previous experience in visual color comparison, while the rest did not have any training whatsoever. On the whole, the results of this test are considered to be a sufficient proof that there exists a visual conception of color differences which can be expressed as a numerical magnitude.

(3) Evaluation of color differences from physical measurements. The problem, as it presents itself now, is to find a mathematical function which will express color differences from physical measurements of colors, and which will agree with

the average visual evaluation of such color differences, as shown in Fig. 3. This problem resolves itself into two parts: first, a numerical color specification based on physical measurements giving three psychophysical parameters for each color; and secondly, a mathematical equation by means of which the six parameters necessary for a specification of two colors could be reduced to a single number. To be of practical value, these measurements must be accurate, reproducible, simple and rapid. The use of Hardy photoelectric spectrophotometer in detecting and measuring small color differences was described by Nutting.[6]

The following methods were employed in measuring the psychophysical color specifications. (1) Spectrophotometric reflectance curves drawn by Hardy spectrophotometer from which dominant wave length, brightness and saturation were computed by the 10mμ method for I.C.I. illuminant C. (2) Disk colorimetry with Munsell disks evaluated in terms of I.C.I. specifications.[7] (3) Judd's wedge colorimeter.[8] (4) Hunter's multipurpose reflectometer.[9]

In evaluating color differences the following procedures were used. (a) Nickerson index of fading.[10] (b) New Nickerson's formula. (c) Judd's equation based on data from spectrophotometry,[11] subtractive colorimetry,[12] and Hunter's reflectometer.

(a) Nickerson's index of fading. A method of evaluating color differences was proposed by Nickerson, based upon Munsell[13] notation of colors in terms of hue, value (lightness), and chroma (saturation). It is assumed for the purposes of the proposal that the total variation in color between two samples may be expressed as a sum of the difference in hue, value, and chroma. Coefficients, based on certain experimental evidence, are assigned to each one of these attributes so as to place them on an equal basis as far as the magnitude of visual perception of color differences is concerned. This sum is designated as the "Index of Fading" and its numerical evaluation is made by the use of the following formula:[14]

---

[6] R. D. Nutting, The detection of small color differences, *Textile Res.* 4, 1934, 323, 323-330; Color tolerance: Its measurement and specification for dyed textiles, *ibid.*, 6, 1935, 104-114.

[7] Dorothy Nickerson, Disk colorimetry; including a comparison of methods for computing tristimulus values for certain disks, *J. Opt. Soc. Amer.*, 25, 1935, 253-257; Use of I.C.I. Tristimulus values in disk colorimetry, *U. S. Dept. Agri.*, May, 1938.

[8] Judd, A subtractive colorimeter for the measurement of small chromaticity differences between surfaces of moderate spectral selectivity of reflectance, *J. Opt. Soc. Amer.*, 26, 1936, 225-230.

[9] R. S. Hunter, Problems in the development of a multi-purpose reflectometer, *J. Opt. Soc. Amer.*, 27, 1937, 225-235; F. Scofield and C. R. Cornthwaite, Measurement of the fading of tints, *Nat. Paint, Varnish, & Lacquer Ass.*, Circular No. 548, 1939.

[10] Nickerson, How can results of fading tests be expressed, *A. S. T. M. Standards on Textile Materials*, Sub-Committee A-7, Committee D-13, October, 1935; The specification of color tolerances, *Textile Res.*, 6, 1936, 505-514.

[11] Judd, Specification of uniform color tolerances for textiles, *Textile Res.*, 9, 1939, 253-263, 292-308. Uniform tolerances for surface-color specification, *J. Opt. Soc. Amer.*, 28, 1938, 52-55.

[12] Judd, Specification of color tolerances at the National Bureau of Standards, this JOURNAL, *supra*, 418-428.

[13] A. H. Munsell, *A Color Notation*, 1936, 1-40.

[14] The evaluation of the 'Index of Fading' by this formula gives a numerical value for I six times as large as the value computed by the first Nickerson's formula, given in the publication referred to. At present formula [1] is recommended by Miss Nickerson for computation of color differences.

$$I = (C/5)\ 2\Delta H + 6\Delta V + 3\Delta C \quad \ldots\ldots\ldots\ldots\ldots\ldots [1]$$

where H stands for hue expressed in 100 steps along the hue circle, V is the value, and C is the chroma in Munsell notations.

In order to evaluate the Index of Fading it is first necessary to express each color in Munsell notations. Using the trichromatic coefficients and the charts prepared by Nickerson, based upon the Hardy-Glenn-Killian data,[15] it is a simple matter to convert I.C.I. specifications into the necessary notations of hue, value, and saturation. The Index of Fading was computed for each of the 10 color-pairs and the ratio

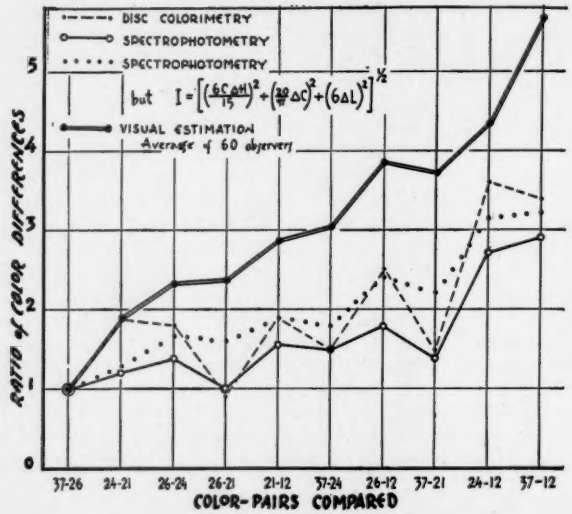


FIG. 4. COMPARISON OF COLOR-DIFFERENCE RATIOS EVALUATED BY NICKERSON'S FORMULA
(60 *Os*)

of color difference was taken to that of a standard pair. These values are represented by circles in Fig. 4.

Munsell notations were also determined from disk colorimetry. The ratios of color-differences were evaluated again for formula [1] and plotted as shown by the broken line in Fig. 4. The discrepancy between the Munsell notations obtained from spectrophotometric curves and disk colorimetry is fairly large and there are probably several independent factors which may account for it.

These ratios of evaluated color differences may now be compared with those obtained by direct visual estimation, also shown in Fig. 4 by a double line. It can be seen that the color differences given by Nickerson's formula are too small

[15] It is expected that these data will be published in the forthcoming Report of the Colorimetry Committee of the Optical Society of America.

by a factor of about two as compared with visual estimations.

(b) Nickerson's new formula. The expression for Index of Fading as given by formula [1] does not represent a metric in a Euclidean space because the terms are not squared and the value of π, as a ratio between changes in hue to changes in chroma is not equal to 3.14. In order to modify Nickerson's formula to a Euclidian form it is necessary to change the relative weights assigned to differences in hue, value, or saturation. If we choose to preserve the same relative weights for the attributes of hue and value, but change chroma so as to conform to the requirements of the Euclidian geometry, we get the following expression for Nickerson's Index of Fading:

$$I = [(c/5 \cdot 2\Delta H)^2 + (6\Delta V)^2 + (20/\pi \cdot \Delta C)^2]^{1/2} \quad \ldots\ldots\ldots\ldots [2]$$

New ratios computed by formula [2] from spectrophotometric data are plotted in Fig. 4 as black dots and show a considerable improvement over the values given by the original Nickerson's formula. It is conceivable that a still better agreement may be attained by a further modification of the relative weights assigned to differences in hue, value and chroma, a probability suggested by Nickerson in her papers.

(c) Judd's equation. Another method of computing color differences was proposed by Judd based upon the use of Maxwell's triangle yielding a uniform chromaticity scale.[16] The equation below gives a mathematical function for evaluating color difference ΔE in terms of trilinear coördinates r, g, and b derived from I.C.I. trichromatic coefficients, and apparent reflectance, A.

$$\Delta E = 600[A^{1/2} (\overline{\Delta r^2} + \overline{\Delta g^2} + \overline{\Delta b^2}) + .005\, \overline{\Delta A^2}/A]^{1/2} \quad \ldots\ldots\ldots [3]$$

The magnitude of the multiplying factor of 600 was chosen so that $\Delta E = 1$ represents a surface-color difference of some significance in industrial or commercial practice. It is about five times larger than the least perceptible difference under good observation conditions. The constant 0.005 depends upon the line of demarcation between the two samples. If the width of the dividing area is imperceptible, this constant should be increased by a factor of 20 or it may be reduced by a factor of 10, if the two surface-colors under comparison are separated by an area of some other color. The purpose of this constant is to make the necessary adjustment in the visual perceptibility of differences in brightness or chromaticity which depend upon the area of separation between the two surfaces under comparison. It is well established that differences in brightness become enhanced in close proximity and are greatly reduced when separated, whereas differences in chromaticity are practically independent of the spatial arrangement.

With the aid of equation [3] and the I.C.I. specifications for 5 green colors, ΔE were evaluated and ratios of color differences computed. These are shown by circles in Fig. 5. Another set of measurements was made on a Subtractive Colorimeter and color differences evaluated from Judd's equation reduced to appropriate terms. These are shown by broken lines in Fig. 5.

Still another way in which Judd's equation can be made the basis for computing color differences is to use Hunter's multi-purpose reflectometer to secure the numerical magnitude of three parameters for specifying a desired color. These

[16] Judd, A Maxwell triangle yielding uniform chromaticity scales, *J. Opt. Soc. Amer.*, 25, 1935, 24-35.

three parameters are the apparent reflectance through Blue, Amber, and Green filters. In terms of these three apparent reflectances, designated respectively by B, A, and G, the expression for color difference ΔE is given by the equation:

$$\Delta E = 730[G^{1/2}(2\overline{\Delta\alpha}^2 + \overline{\Delta\beta}^2) + \overline{\Delta G/300G}^2]^{1/2} \ldots\ldots\ldots\ldots [4]$$

where $\alpha = (A - G)/(3B + 2A + 5G)$ and $\beta = (G - B)/(3B + 2A + 5G)$. The green filter in the reflectometer was selected so as to give in combination with the spectral sensitivity of the photoelectric cell a response closely approximating the luminosity function for a normal eye. It follows, therefore, that the numerical value of luminous apparent reflectance, A, used in equation [3] is identical with the

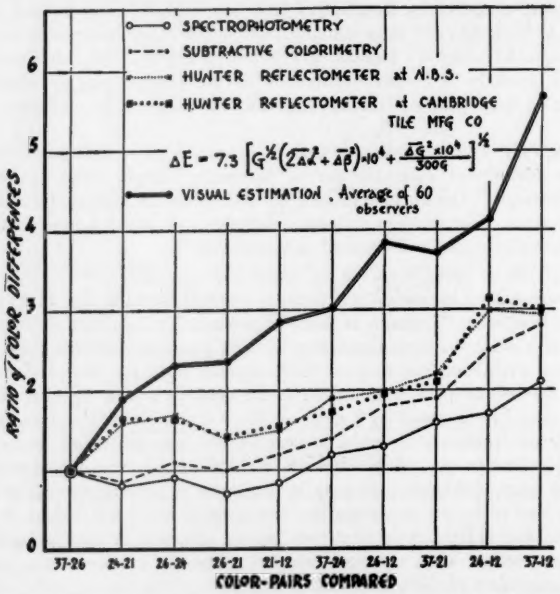


FIG. 5. COMPARISON OF COLOR-DIFFERENCE RATIOS EVALUATED BY JUDD'S EQUATION
(60 *Os*)

apparent reflectance, G, as measured directly on Hunter's reflectometer. The constants in equation [4] were adjusted to give the same numerical magnitude for unit of color difference as in the original Judd's equation [3].

For the purpose of comparison, all the values of color differences were reduced to the arbitrarily chosen unit of color differences between titles 37 and 26, which came out to be equal to 5.81 units. A sample calculation of color differences between tiles 37 and 26 is shown in Table II. Following a classical tradition of naming a unit after the man who was primarily responsible for developing the basic principles upon which such measurements can be performed, it is suggested here to designate such unit of color difference by the name "judd."

The ratio of color differences computed from Judd's Equation on the basis of measurements made by the three physical instruments described above are plotted in Fig. 5. The examination of all the graphs will show that the methods of color-difference evaluation proposed by Nickerson and Judd are far from giving a perfect agreement with the average estimations made visually.

The hope for a better agreement lies in the proper adjustment of the formula so as to modify the relative weights assigned to differences in hue, lightness, and saturation. Such work is now in progress. It must be remembered that the individual estimates of color differences may vary by as large a factor as ten and it is very significant that at least the ranking of color differences in visual observations and mathematical computations based on physical measurements follows practically an identical order. From the point of view of agreement, Nickerson's new formula gives the best approximation to visual estimations. From the standpoint of perform-

TABLE II

SAMPLE CALCULATION OF COLOR-DIFFERENCES BETWEEN TILES 37 AND 26

Evaluation of color-differences: $\Delta E = 7.3[G^{1/2}(2\Delta\alpha^2 + \Delta\beta^2) \cdot 10^4 + (\Delta G^2 \cdot 10^4)/300G]^{1/2}$

No.	1	2 Reflectance		3	4	5	6	7	8	9	10	11	12	13
	Blue	Amber	Green	A−G $\times 10^2$	G−B $\times 10^2$	3B +2A+5G	α Col. 4/6	β Col. 5/6	G Average	ΔG $\times 10^2$	ΔL	ΔC	ΔE judds
26	.460	.446	.514	−6.8	5.4	4.842	−1.404	1.115	.561	9.4	5.36	2.24	5.81
37	.535	.537	.608	−7.1	7.3	5.719	−1.241	1.296					

ing physical measurements of color, Hunter's multipurpose reflectometer offers the most rapid, sensitive and sufficiently reproducible means of determining the three color parameters. Under these conditions it was felt that it would be a worthwhile attempt to adopt this instrument in conjunction with Judd's equation to the solution of several practical problems dealing with industrial color tolerances.

(4) Types of color space. The fact that any color can be completely specified by three parameters suggests a possibility to represent any color as a point in space, whose three coördinates give directly the values of the three color attributes. Colors, therefore, may be considered to correspond to a manifold of points defined by a selected group of three coördinates. Interpreted geometrically, such manifold constitutes a color space.

In a color system established by Munsell, the central vertical axis corresponds to a series of achromatic colors beginning at the bottom with a perfect black and ending at the top with a perfect white. Along the planes perpendicular to the neutral axis are located all the colors of one value (lightness). Colors of different chroma (saturation) are represented in this chromaticity plane by their distance from the axis. Any point on the circle, drawn around the achromatic axis as a center, has the same saturation which increases as the radius of the circle is made larger. Hue is associated with the angular distance from an arbitrary point on the "hue circle." The colors in Munsell space are so arranged that equal linear distances along the vertical or radial axes, or equal angular displacements along the hue circle represent equal steps in sensation *within each color attribute.* This distribution of color variations may be associated with the property of the space itself. Whenever colors are set in order spatially in accordance with visual perception, as the only basis for classification, such an arrangement constitutes what might be called a *psychological* color space. Of this, Munsell's is a characteristic example.

Let us consider two points representing two different colors in this Munsell color space. The linear distance between these two points we may attempt to associate with the magnitude of color difference between them. Such direct evaluation, however, will not agree with visual perception of color differences when compared with another set of two points whose distance of separation may be the same. The reason for this is that one step in hue is not equal in perceptibility to one step in value or chroma. In fact, Nickerson's formula is an attempt to introduce a certain multiplying factor for two of the coördinates so that numerically all three color attributes are placed on the basis of equal perceptibility.

It would be possible also to arrange in space different colors according to their apparent reflectances as determined physically from the spectrophotometric curves. In this system, the word "color" could be used, if at all, only in a very limited sense devoid of all implications in reference to human vision. Instead of a luminosity curve as the basic factor in color perception we could substitute some function characteristic of spectral sensitivity of a certain photoelectric cell. Such a spatial arrangement based entirely upon physical measurements may be referred to as a *physical* color space.

By combining the concepts of human eye sensitivity together with those of physical measurements we arrive at a *psychophysical* color space. The I.C.I. specifications, indeed, allow us to determine a triad of numbers which assign unique position to any given color in such a space. It is valuable in so far as the position of every point represents a definite color based upon physical measurements and expressed in terms of visual perception. It disregards entirely, however, the magnitude of color differences as related to the linear distances.

For the purpose of evaluating color differences we must evolve, what might be called, a *homogeneous* and *isotropic* psychophysical color space. Such a space would satisfy the following requirements. (1) Equal linear distances between points representing any two color-pairs correspond to equal sensations of color differences. (2) A larger linear distance between two points corresponds to a greater perceptible color difference. (3) The ratio of linear distances between two pairs of points agrees numerically with the visual estimation of color differences of these two color-pairs estimated by the ratio method.

Whether such an homogeneous and isotropic color space can be represented in three dimensions, or whether it will have an Euclidean metric cannot be as yet definitely answered. Indeed, there are some like Boring,[17] who propose to extend the number of independent variables from three to five. The possibility of using Euclidean geometry in a color domain is also questioned by Silberstein.[18] There is no question about the desirability of limiting color specification to a space of three dimensions and retaining in it the simple relationships of Euclidean properties.

The investigation on industrial color tolerances described in this paper was, in fact, an attempt to deal with a color space as defined by Judd's equation, in terms of color measurement data of Hunter's reflectometer. The expression for ΔE given in Equation [4] becomes, therefore, a metric in a color space referred to three rectangular axes of α, β, and G. The chromaticity plane is a plane of constant G,

[17] E. G. Boring, Isochromatic contours, this JOURNAL, 49, 1937, 130-134.
[18] L. Silberstein, Investigations on the intrinsic properties of the color domain, *J. Opt. Soc. Amer.*, 28, 1938, 63-85.

where various values of α and β determine the hue and saturation of a color. The coördinates α_1, β_1, will determine the position of a point in the chromaticity plane and G_1 is plotted along a perpendicular erected from the point α_1, β_1. In a chromaticity plane where 1 $\Delta\alpha \times 10^3$ plotted graphically as 100 cm., the proper factors to be used for ΔG so that a linear distance between two terminal points will represent a color difference equivalent to 1-judd are shown in Fig. 6.

The geometrical method of evaluating color differences is applicable to problems

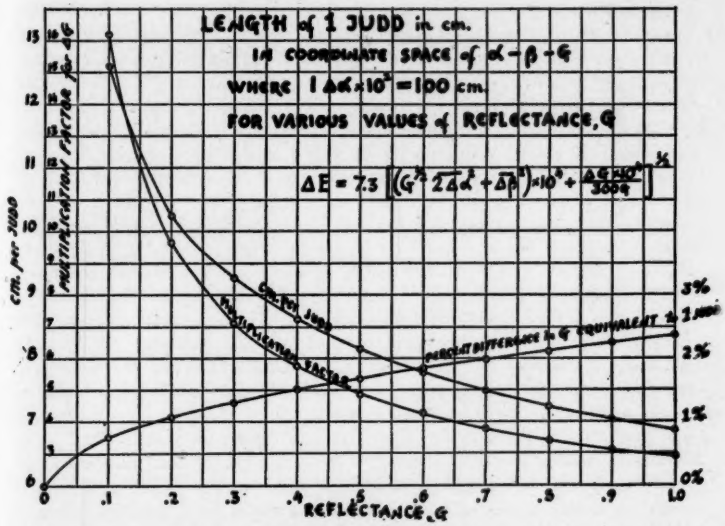

LENGTH of 1 JUDD in cm.
IN COORDINATE SPACE of α-β-G
WHERE 1 Δα×10³ = 100 cm.
FOR VARIOUS VALUES of REFLECTANCE, G
ΔE = 7.3 [(G²Δα² + Δβ²)×10³ + ΔG²10²/3G]^½
CM. PER JUDD
MULTIPLICATION FACTOR for ΔG
REFLECTANCE, G
PERCENT DIFFERENCE in G EQUIVALENT to 1 JUDD
3%
2%
1%
0%

FIG. 6. TRANSFORMATION OF COÖRDINATES FOR NUMERICAL EVALUATION OF COLOR-DIFFERENCES

where color differences are not large or where reflectances are high as can be seen from Fig. 6. This satisfies the conditions found in most of the problems on industrial color tolerances. The two samples, E_1 and E_2 are measured in Hunter's reflectometer and their reflectances in the spectral transmission regions of blue, amber, and green filters are determined. From these α_1, α_2, β_1, and β_2 are determined and plotted on the chromaticity plane. ΔG is then computed and multiplied by a factor, m, as given in the graph of Fig. 6, corresponding to an average reflectance $G = (G_1 + G_2)/2$. From a point (α_1, β_1) which represents a color with greater G, a perpendicular is erected equal to $m\Delta G$. The distance between the terminal point of the perpendicular to the point (α_2, β_2) in the chromaticity plane is then equal to color difference in judds. The linear distance in cm. equivalent to 1-judd for different average reflectances can be found from the graph in Fig. 6.

This geometrical method of evaluating color differences is particularly well adapted to problems where color differences are to be evaluated between pairs of every pos-

sible combination of many colored samples. The procedure, then, involves a construction of a color space model in which every color is represented by a terminal point of a steel needle set to a proper height above a chromaticity plane. Multiplication factor, m, is determined by the average reflectance of all colors. A special measuring stick is graduated directly in judds, the length of a judd being determined again by average reflectance. Such a physical model gives at a glance the relative distribution of colors in a limited domain of a color space. The linear distances between terminals of any two colors within this domain are proportional to the sense perception of color differences.

(5) Tolerance boards. From the standpoint of consumer's acceptance of maximum color variation, the most direct method of finding out the proper color tolerance, would be to present for their approval properly selected samples in which the range of maximum color variation would vary within predetermined limits. The consumers are then to be asked what maximum color tolerance they would consider acceptable. Each exhibit is to be prepared in such a form as to simulate the actual conditions under which these color variations are to be seen when the materials or goods are in actual use.

In certain instances the variation in color may be limited to two extremities within a single attribute of color, such for example as the variation in brightness of roasted coffee beans. The samples could be arranged giving a standard shade and different gradations from light to dark. After a careful statistical analysis certain material standard tolerances could be established giving the color tolerances for maximum permissible variations from standard toward a lighter or darker shade. Such material standards of color tolerances are in daily industrial and commercial use. Very frequently they are set more or less arbitrarily without any resort to a direct statistical study of consumer's reactions.

In the field of a glazed ceramic wall tile the variations are usually found in all the three attributes of color. The presentation of three or four tiles would be entirely inadequate and it was therefore decided, in order to carry out this direct method of finding consumer color tolerances, to use at least nine tiles arranged in the form of a square with three tiles on each side. These "tolerance boards" were made of carefully selected tiles which gave a uniform color variation within a maximum color difference which was selected to be: one, two, three, four, and five judds. Whereas the color differences between two adjacent tiles A-B, B-C, C-D, etc., up to H-I were equal, they progressively increased from tile A up to tile I, to wit:

$$\text{A-B} = \text{B-C} = \text{D-E} \ldots\ldots\ldots\ldots\ldots\ldots \text{H-I}$$
$$\text{A-B} = (\text{A-C})/2 = (\text{A-D})/3 = \ldots (\text{A-I})/8$$

If the maximum color variation between two extreme tiles of a tolerance board is, say, 4 judds, the variation from tile to tile is to be only ½ judd, the interval variation being one eighth of the maximum color variation. The properly selected tiles are mounted on Masonite boards in the following relative possitions.

B	C	D
I	A	E
H	G	F

It is possible to arrange these 9 tiles in such order that the appearance of the board

from the standpoint of color variability is either enhanced or weakened. The above arrangement was selected because it gave a relatively true representation of color variations as a whole and it was simple in showing an orderly and gradual change in color from tile to tile.

According to specifications described above, the preparation of tolerance boards calls for tiles with uniform interval of color variation so that a group of nine tiles would give a maximum color range of one, two, and up to five judds. Although,

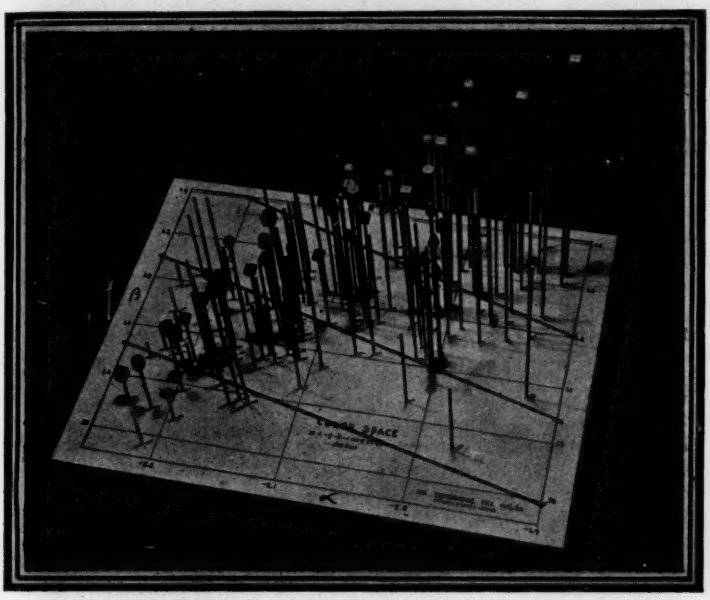


FIG. 7. A MODEL REPRESENTING AN ARRANGEMENT OF COLORS BY POINTS IN A COLOR SPACE OF α-β-G

Each group of nine selected tiles is shown by differently shaped figures attached to the terminal points. The following shapes designate tiles that were selected for different tolerances: cones—1 judd; pyramids—2 judds; balls—3 judds; cylinders—4 judds; and cubes—5 judds.

at first, the possibility of making such selection did not appear to be practicable, later on, the development of the geometrical method of representing colors and evaluating color differences, made the solution of this problem feasible.

About 200 tiles of a desired color were picked at random from different production runs and each assigned a number. The color of each tile was measured on the reflectometer and the chromaticity components α and β were computed. These points were then plotted on the α-β plane and designated by their respective numbers. Pins were then inserted into the chromaticity plane to represent the reflectance of every tile. A "yardstick" for measuring color difference in judds was made in

accordance with the average reflectance of the whole group of tiles. It was surprising to find that within a group of two hundred tiles there were very few tiles that showed identity in all three color attributes. By visual inspection nine consecutive points, all lying approximately equidistant along a straight line, were selected to give the necessary maximum color variation for each tolerance board. A photograph of an actual model built for selecting tiles is shown in Fig. 7.

The surprising thing about this procedure was the ease with which 45 tiles for five different tolerance boards could be selected. It is not to be expected that an exact equality in the spacing could be achieved. The accuracy of selection was sufficiently high that, when numbered tiles were actually laid out in a row, the color differences from tile to tile appeared to be of the same magnitude. In three or four instances visual observations indicated a discrepancy in the gradation of color differences. These were later on traced to an error in computations or the geometrical construction, determining the location of the terminal point for that particular

TABLE III

COLOR TOLERANCE IN JUDDS
(46 *Os*)

	Glaze number		
	74	80	82
Most desirable	1.3	2.2	1.9
Max. acceptable	2.4	2.8	2.8

color. By the use of such color space models, tiles were selected to make 5 tolerance boards for three different colors.

The availability of tolerance boards offers the manufacturer a very specific way of finding out what the consumer wants and will tolerate as far as the color variation in the product is concerned. A set of 5 tolerance boards, numbered at random, is exhibited under proper illumination. A selected group of people is asked to write down the number of a tolerance board which shows either the most desirable or the maximum acceptable color variation. The Color Variation Ballot for three different colors reads as follows:

Granting that variation in color from tile to tile is one of the distinctive features of the product, and assuming that the factory is able to maintain and control the variations as shown on the board.

(1) Which in your opinion represents the *most desirable* color variation?

74	80	82
Dusty Pink	Robins Egg	Cafe au Lait
Board no. ( )	( )	( )

(2) Which board in your opinion represents the *maximum* possible color variation, which would be acceptable to you as a consumer?

Board no. ( )	( )	( )

The results of voting by 46 *Os* are given in Table III. The average tolerances, T_a for different colors were computed by the following formula:

$$T_a = (\Sigma^5_{K=1} \cdot K \cdot V_K)/(\Sigma^5_{K=1} \cdot V_K), \text{ where}$$

V_K is the number of votes cast for a tolerance board with a maximum color difference of K judds.

These preliminary results indicate that there is a very good consistency in correlating visual estimations of color differences with the numerical evaluations of these differences by Judd's equation. There seems to be an indication, however, that color differences in the spectral range of red corresponding to glaze 74 are too small when computed from Judd's equation. Tolerance boards can be used, therefore, also as a check on the homogeneity of different domains within a color space.

It would be also possible to select tolerance boards within any given domain so that colors differ by the same amount in hue, brightness, or saturation. These boards could be intercompared and conclusions drawn whether a variation in one attribute is equivalent to the same numerical variation in another color attribute. This will constitute a test for isotropy of any color domain.

(6) Numerical values of color tolerances. The procedure outlined above allows the manufacturer to obtain valuable information regarding either desirable or maximum permissible color variations. The discrimination between these two types of color variations does not necessarily exist in all commercial products. Whenever such distinction can be made it will lead to a better satisfaction of consumers' wants. It is of interest to record a spread of approximately one judd as voiced by a group of voters in regard to color variations in ceramic wall tiles. This fairly narrow limit of color variability may appear to be more severe than the actual use of the product calls for, yet such was the response of the observers. As pointed out by Judd, "the manufacturer would be only too glad to use larger tolerances, but the purchaser in many cases either does not understand the necessary connection between size of color tolerance and cost per accepted unit, or else he uses the color tolerance as a loophole through which to escape an unfavorable contract when there is no other way to escape it."[39] If the size of color tolerance could be definitely specified and accepted, it would serve well to guard the interests of both the manufacturer and the consumer.

The numerical magnitude of color tolerance must depend upon the specific usage or utility of the goods where color differences serve as indicators of variations in certain desired properties. It is, therefore, futile to expect that a single color tolerance could ever be established. Within a specific industry, or in related fields, definite standards of color tolerances would be most desirable.

It might be well to classify different groups of color use which may affect the magnitude of color tolerances. (1) Color in reproduction of objects as they occur in nature. Color in art, printing or photography. (2) Color as indication of quality, chemical composition or desired properties. Color in agricultural products, colorimetric analysis or pH determinations. (3) Color as a means of identification. Color of signal glasses, trade marks, various color designations as in army or navy. (4) Color standards as used for references in various systems of color classification. Munsell color atlas, Ostwald sytem, Merz and Paul dictionary of color. (5) Color as an element of fashion or in compliance with conventional custom, style or usage. (6) Color as related to other colors to produce pleasing aesthetic effects or harmonious combinations.

Numerical values of color tolerances may be of course different within each group and for each article depending upon the condition of use, manufacture, and

[39] Judd, Letter to the editor, *Bull. Amer. Cer. Soc.*, 17, 1938, 379-380.

price. Although in some industries the general idea of a proper working color tolerance may be fairly well established by the use of material standards, this situation is far from universal. It is reasonable, however, to expect that the size of color tolerances will be generally much smaller in one group than in another. For example, we can hardly expect that a color match in a trade mark requires the same degree of reproducibility as would be desirable to have in material color standards. Speaking in terms of numerical color differences, 3 or 4 judds might be acceptable in the first instance and a fraction of a judd would be desirable in the second.

Sometimes it may be even necessary to establish different sizes of color tolerance for different spectral regions. In color photography the appearance of flesh would certainly require a smaller departure from the actual color, than the color of a dress. In other instances a larger color deviation from a standard in a direction of one hue may not be as objectionable as an equal color deviation toward another hue. As an illustration we may cite the example of butter standards, where a shift in hue from yellow to orange could be tolerated to a greater degree than a shift toward green. These two instances are mentioned to show that every product requires a careful study before color tolerances can be established. In both of these cases there is no reason why a numerical specification of color tolerances could not be set up to take care even of these exceptional problems. This can be done by varying the size of color tolerance for different spectral regions or for different directions in color space.

(7) Control of color uniformity. There are usually several factors which determine the color of a finished product. In mass-production methods these factors depend upon certain physical and chemical tolerances which are set so that the maximum color variation will not exceed a certain limit. It must be obvious, therefore, that a *consistent* color uniformity of manufactured products depends entirely upon the tolerances maintained in these physical and chemical factors. The use of the adjective "consistent" is necessary because in certain instances it is possible to achieve a uniformity in color if it so happens that the variations in any two factors are in such a direction that they neutralize each other.

It follows, therefore, that for production of articles within a certain color variability it is necessary, first, to be able to measure numerically each one of the physical and chemical agents and, second, to determine the effect of variation of these agents upon the final color. If more than one agent is involved it will be necessary to introduce a measured variation into a single agent, maintaining all the others constant. The results of such tests will supply the information in regard to the maximum permissible variation in physical and chemical agents to produce a variation in color not exceeding a certain color tolerance limit. As an approximation we may assume that the maximum color deviation from a standard is given by the sum of maximum color deviations of each one of the independent agents.

In the manufacture of ceramic tiles we will consider the effect of only two factors upon the color of a final product: (1) Thickness, a, of the fired glaze; and (2) temperature, t, at which it was fired. If ΔE represents a color difference between a standard and a sample tile because of a difference in the thickness of the two glazes and temperatures at which they were fired, then:

$$\Delta E = \Delta a \, (\delta E/\delta a)_t + \Delta t \, (\delta E/\delta t)_a, \quad \ldots\ldots\ldots\ldots\ldots\ldots [5]$$

where, $\delta E/\delta a$ is a partial derivative of color with respect to glaze thickness, $\delta E/\delta t$

is a partial derivative of color with respect to temperature of firing, while Δa and Δt are measured differences in thickness and temperature respectively.

It can be observed from the above equation that for a maximum desirable value of ΔE, the factors under the manufacturing control are determined entirely by Δa and Δt, since $\delta E/\delta a$ and $\delta E/\delta t$ are the intrinsic properties of the glaze itself. It is only when the numerical values of ΔE, and $\delta E/\delta a$ and $\delta E/\delta t$ are known, that we

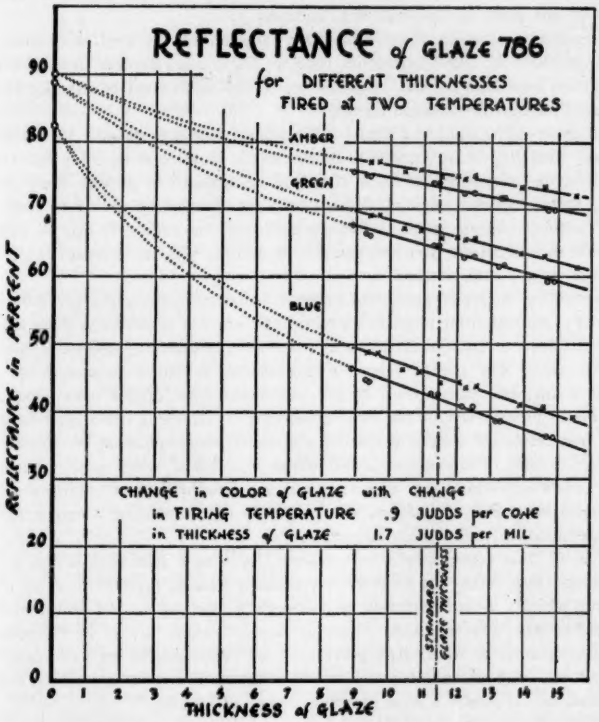


FIG. 8. REFLECTANCES OF TILES OF VARIOUS GLAZE-THICKNESSES FIRED AT TWO DIFFERENT TEMPERATURES AND MEASURED ON HUNTER'S REFLECTOMETER

can set up a tolerance for Δa and Δt. If, from the standpoint of industrial control, Δa and Δt are too large, then we have to be content with a larger value of ΔE or change the basic properties of the glaze under consideration.

The use of the Color-Uniformity Control equation [5] provides the industry with specific information upon which to establish a rigorous control of color uniformity.

The determination of rates of color variation as a function of thickness or temperature is made by the application of glaze in gradually increasing thicknesses to

two sets of, say, 10 tiles. Each set is fired at two different temperatures. The color of each tile is then carefully measured on Hunter's Reflectometer and the reflectances B, A, and G plotted as shown in Fig. 8. From this graph computations are made to determine $(\delta E/\delta a)_t$, as so many judds per mil of glaze thickness, and $(\delta E/\delta t)_t$, as so many judds per degree Fahrenheit or per cone. With these values at hand tolerance limits for Δa and Δt can be set up so that ΔE would not exceed a definite amount. These values are also used to determine the degree of relative stability of one glaze as compared with another.

A numerical evaluation of rates of color variation can be applied to many industrial problems of color uniformity control. These rates become definite criteria for a general expression of color permanency of any material as affected by various measurable physical or chemical agents.

Conclusions—The results of visual tests indicate that it is possible to establish a numerical scale for visual perception of magnitude in relation to color differences. A considerable amount of additional data will be required to provide the information for various spectral regions. Persons who are definitely planning to enter this field of investigation are asked to communicate with the author. It may be possible to provide them with the necessary samples of ceramic tiles for estimations of color differences by the ratio method.

Although the degree of agreement between visual estimations of color differences and those computed from physical measurements are not satisfactory, their application to problems of consumers' acceptance and tolerance control proved of practical industrial value. The use of Hunter's reflectometer for these measurements gave fairly reproducible results within $\pm$ 0.1 judd under most careful operational technique. Further improvements are needed, however, to achieve a better reproducibility of this instrument for routine industrial use under plant operating conditions. The reduction of color measurements to differences in terms of visual perception by the use of Judd's equation is still a very tedious and lengthy process requiring at least ten minutes for a pair of samples. The possibility of constructing a nomograph for such computations is now being investigated.

All in all, this paper should be considered only as a trial step in the field of color engineering. What the industry is ultimately looking forward to is an instrument into which two surfaces could be placed for a direct numerical measurement of color differences. The indicating scales of this instrument are to be calibrated in terms proportional to visual perception of color differences in each of their three attributes and also as an integrated single number. The development of such an instrument would render a most valuable service to industry.[20]

[20] Acknowledgment is made to the following, for only by their generous coöperation has this paper been possible: Dr. Lyman J. Briggs, Dr. K. S. Gibson, Mr. R. S. Hunter, Dr. Deane B. Judd of the National Bureau of Standards; Miss Dorothy Nickerson, U. S. Department of Agriculture; Dr. R. C. Gowdy, University of Cincinnati; and particularly Mr. Charles H. Burchenal, Cambridge Tile Manufacturing Company.

APPARATUS

VISUAL-TACTUAL ILLUSIONS FROM A CRYSTAL BALL

By WALTER R. MILES, Yale University

One hesitates to say what has and what has not been seen in crystals. There are, however, some rather matter-of-fact illusions to be experienced from holding a small crystal ball in the hand, which so far as the writer knows, have not hitherto been described and would seem to merit notice in our psychological literature. For several years past these visual-tactual illusions have been demonstrated to small classes, and it has been found that the majority of participants profess to get the effects quite readily. Undoubtedly suggestion has been present in some cases, but the illusions satisfy the criterion of persistence with repeated trials and in general tend to grow more conspicuous. They depend primarily on the unique optical properties of the sphere which may be briefly reviewed before describing the illusory experiences themselves.

The crystal ball is a poor lens because it has the maximum of spherical aberration. Parallel beams of light which strike the sphere nearly normally to its surface, that is, in a very small region around a central axis, are focussed fairly sharply, but the rays striking the outer zones are refracted more and cross earlier, forming a caustic surface that gives marked image distortion.[1] For example, if a solid glass sphere 70 mm. in diameter is placed with its center 1 meter from a bright object, such as a reading lamp surrounded by an ordinary parabolic reflector, a fairly sharp image of the circular contour of the reflector will be formed on a screen placed about 17 mm. behind the sphere. To an *O* in line with the optic axis and 50 cm. (arm's length) from the sphere, with no screen interposed, the image on close inspection will appear to be turned wrong-side out. The lamp itself seems to be in front, while the surrounding reflector shows as a plane, or as a convex surface behind it. Furthermore, the image of the reading lamp will at first appear to be at or on the surface of the glass sphere, but if a mark which can serve for visual fixation is placed on the sphere on *O*'s side and next the image, the latter will then be seen not on the glass surface but standing in front of the mark at a distance of about 13 mm. A pencil or other small object may be inserted in the space between the fixation-mark and the newly identified position of the image seen in binocular vision. As the image is made to move at right angles to the line of vision it approaches the sphere along a curved or caustic line and meets it near the apparent circumference. Images formed by the crystal and the path of an image moving across the face of the crystal both suffer caustic line distortion. If, however, the crystal sphere is free from all marks that can serve as fixation-points, *O* is not aware of this distortion and has the impression that the image is in spatial register with the surface of the sphere in the same way that a map fits the surface of a globe. The image-field and its movement are accepted as characteristic of the translucent sphere as a self-contained optical instrument. This perceptual misinterpretation of

[1] See Helmholtz's *Physiological Optics*, 1, 1924, 66.

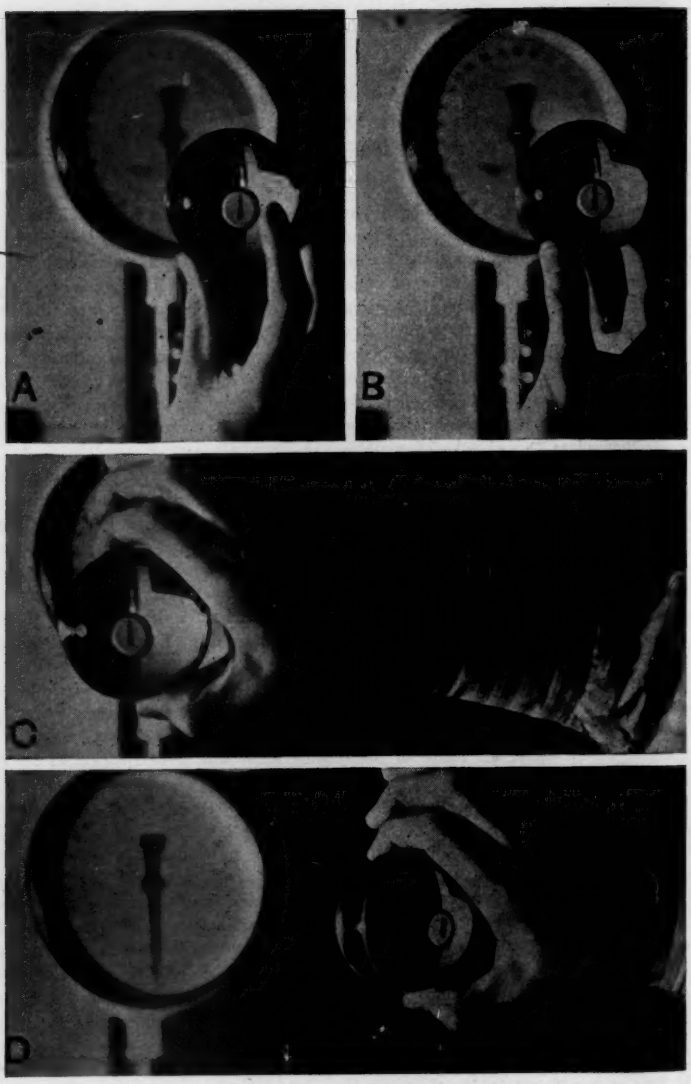


FIG. 1. VISUAL-TACTUAL ILLUSIONS FROM A CRYSTAL BALL

position is in itself a visual illusion that escapes notice and supplies a datum of experience around which other illusions involving vision and touch may occur.

No very special arrangements are requisite for observing visual-tactual illusions in connection with the crystal ball. It is desirable but not necessary to use a complex object field so that the visible convex hemisphere will appear to be well filled with image forms without too much transparent space between them. *O* should grasp the crystal lightly, placing his fingers in such a manner as not to cover the central portion on either side. It is important that the crystal should be spherical without distorting streaks in the glass, and that its surface be clean.

Two forms of the visual-tactual illusion may be distinguished: (1) an illusion of rest, and (2) an illusion of motion. An attempt has been made in Fig. 1 to illustrate both forms. A and B show terminal views for the illusion of immobility of the sphere, while C and D give respectively the beginning and end of an illusion of counter-clockwise turning of the sphere on the axis formed by the thumb and forefinger.

To experience the illusion of rest, *O* supports the crystal on his fingertips and revolves it about its vertical axis without displacing the sphere laterally or vertically. If the viewpoint and object field are both stationary and the ball is not displaced during revolution, the image seen at the surface of the crystal will of course appear unchanged, while *O* alternately executes small right and left rotary movements of his hand. On intently watching the unchanging image, during a few such movements of the hand, one sees the crystal suddenly appear curiously stationary in space, while the fingers are merely gliding around its lower hemisphere. The tactual impressions from the fingertips undergo a temporary sensory eclipse which is in part probably due to adaptation phenomena. The lively sense of contact may disappear and remain absent as long as the sphere "stands still," or a more active tactual impression of a positive nature may come to the foreground of experience. This may be described as a frictionless gliding of the skin areas that are in contact with the glass, as if it were lubricated with an unseen film of oil. The tactual element in the illusion is somewhat transitory and characteristically difficult to analyze. The visual component is very clear and convincing and can usually endure small degrees of displacement in the image without disrupting it. This is especially true when *O* watches the ball while it is being rotated by someone else.

The illusion of rest is fully as striking when the sphere is held with the hand and arm above it and the observer looks obliquely down at the image while rotating the glass. In this position a larger surface of the hand is usually in contact with the ball, but the fingers can be so placed as to avoid covering its visible center. When held upside down it can be rotated well and for a prolonged period without much displacement, and *O*s quickly gain the impression that the hand is gliding over a beautifully smooth stationary sphere such as might decorate the top of a pedestal on a stairway railing. The sphere of course cannot actually be rested on a pedestal during the illusion without producing the visual distraction incident to the magnification of the supporting surface.

The visual-tactual illusion of motion depends on image displacement. If a crystal ball is rolled along the table at the height of *O*'s eyes and at right angles to the line of regard, the regular movement of the image field seen on the surface of the sphere will give the crystal the appearance of spinning about its vertical axis. This visual illusion of motion is equally convincing when the ball, held between the thumb and

forefinger, as in illustrations C and D, is moved laterally. When the hand is moved to the right, as from C to D, the crystal appears to revolve counter-clockwise, whereas it presents a clockwise motion when the hand moves from the right back to the left. Alternate movements of the hand provide the requisite conditions for an irresistible visual illusion of alternate partial revolutions of the sphere. Here again the tactual component provided by the stimulus of the crystal against the fingertips must either subside and take an inconspicuous place in experience or accommodate itself to what is visually apparent and dominating. No great stretch of the imagination is necessary to have faint tactual impressions that the crystal is turning, first one way and then the other, between the fingertips. The illusion is interrupted at the terminus of each swing, but quickly reëstablishes itself when motion starts in the opposite direction. The positive tactual component seemed less stable than the visual element in the illusion, but this is probably in part a result of visually dominated attention.

If the crystal is held as in C but is raised and lowered, it gives the appearance of revolving on a horizontal axis while the fingers glide along its circumference. Here the tactual illusion is rather similar to that described in connection with the illusion of rest.[2]

Illusions are interesting because they present a pseudo-pathology of perception and give oblique glimpses into the structure of experience. They are also useful as tests of rigidity in experience since they demand a high degree of adaptability on the part of *O*. Rigid perceptual patterning operates to extinguish the transitory faint cues out of which a good many illusory psychological structures are built. Therefore, mental attitude, affect and experience are important factors in the person to whom one attempts to introduce a sensory illusion. In not a few cases the attempt fails. Some people appear to be frightened by such experiences and resent any trespass of this character on their established reality domains.

The visual-tactual illusions of the crystal ball, like the size-weight and other illusions, demonstrate the strong tendency of vision to dominate perceptual situations. When there is a conflict between visual perception and the impressions mediated by other sensory modalities, vision assumes or wins the privilege of demanding that its impressions be heeded. Figuratively speaking, vision talks the other sensory impressions out of court or into tacit agreement with its convictions. Even Aristotle's illusion is frightened away if vision sticks in a 'finger.'

---

[2] Other interesting effects of contradictory movement can be gotten from watching the image of one's feet and the floor in a crystal which is being carried in walking forward or backward or in pivoting, and also from placing the crystal on a support and walking around it. A group of four glass spheres attached together as a pyramid, when moved laterally, provides the conditions for a rather pretty visual illusion in which all the balls turn in the same direction in defiance of mechanical forces rigidly holding them together.

NOTES AND DISCUSSIONS

The Harvard *Explorations in Personality*

Explorations in Personality has recently been published from the Harvard Psychological Clinic. The authors of this striking and rather bewildering book have been engaged in a pretentious coöperative research program with three aims: (1) to formulate a dynamic theory of personality; (2) to construct new concepts and experiments for isolating the variables of personality; and (3), as the goal of everything else, to construct psychological portraits or "abstract biographies" of individual subjects. This book is the first report of the undertaking. It provides a rather impressionistic but presumably undistorted picture of the research done in the past five years at the Clinic. Woven into this account is a more extended but somewhat jumbled exposition of the conceptual thinking which was partly the inspiration of the research and partly its outgrowth.

The Harvard Clinic was founded and partially endowed by Morton Prince. Under Dr. Henry A. Murray, its director since 1928, both Prince's medical (clinical) and his psychological affiliations have been perpetuated. Dr. Murray's unusual professional career, leading through the Columbia medical school (M.D. in 1919), the biochemical laboratories of Cambridge University (Ph.D. in 1928), and Jung's institute at Zurich, has now carried him farther into psychology proper than Prince ever got. So far as we can judge from this book his interests at the present time converge squarely on personality studies of the normal individual. But Murray has brought with him to psychology the outlook of a practicing psychoanalyst. This fact makes him unique among those who hold a major position in psychology in a leading American university. It also explains how it comes about that *Explorations in Personality* is by far the most comprehensive attempt to bring Freudian psychology and experimental psychology into line with each other.

The title page of the book bears witness to the unusual spirit of coöperation which has prevailed at the clinic. After the subtitle: "A clinical and experimental study of fifty men of college age," we find authorship of the book credited to "The Workers at the Harvard Psychological Clinic" (in large type) followed by Henry A. Murray (smaller type with capital letters) and by 27 other names (same size type but no capitals), including 2 M.D.'s and 14 Ph.D.'s. Four of the group, we are told (p. 722), had been practicing psychoanalysts. From the separate chapter headings of the eight long chapters in the book it appears that Murray wrote five entirely, shared the authorship of a sixth, and contributed a good-sized introduction to the seventh. The remaining chapter on Procedures, exceeding 200 pages in length, is divided into 25 sections, of which Murray wrote four and must have had much to say about what went into the others.

The bulk of the writing, then, is Murray's. Though the book lacks the unity usually found in the report of an experiment, the consonance of the book's *tone* testifies that one man inspired and to a degree integrated the activities of many coöperating psychologists. Murray has been generous under these circumstances in not publishing more conspicuously under his own name. From his preface, and at frequent intervals later, we glimpse a vast amount of experimental activity going on behind the scenes which never saw fruition in the published report. We can

also sniff, all the way along, the odors that arose from a tremendous broth of discussion stewing for years. Murray's own list of those who threw something into the pot reads: "poets, physicists, sociologists, anthropologists, criminologists, physicians; democrats, fascists, communists, anarchists; Jews, Protestants, Agnostics, Atheists; pluralists, monists, solipsists; behaviorists, configurationists, dynamicists, psychoanalysts; Freudians, Jungians, Rankians, Adlerians, Lewinians, and Allportians. To the fact that we never found a language suitable to all (sic), that some of the experimenters entertained reservations to the last, the reader can ascribe some of the annoyance or pleasure he may experience when here and there throughout the book he encounters varieties of terminology or theory" (p. xi).

Our comment is that, far from the result which this quotation might lead one to expect, all the exposition of fundamental theory in this book is written by one man. It is Murray's distillation of the broth. Those who "entertained reservations" are not explicitly heard from. In the book as it is published whatever disagreements existed are neither set forth nor argued out. There is no weighing of alternative views and no cross-fire of discussion between chapters. This, doubtless for good practical reasons, is all that could be expected in the summary of so extensive and so daring an undertaking, one that Murray himself candidly calls "excessively ambitious." More serious, however, is the fact that the experimental program is not brought sufficiently into line with what has been done outside Harvard, a fault which taints all that the clinic did with idiosyncrasy, and worse. The reasons that lie back of this we may easily infer from the temper of a passage which we shall presently quote in full wherein the standard methods adopted for the solution of such problems are dubbed "statistical, trivial and artificial." One clear fact about the clinic group is that it is working in semi-detachment from the main traditions of differential psychology. They have derived little of positive value from this field and have so far contributed little to it. In the 1939 edition of Hildreth's bibliography listing 4000 items in the field of Mental Tests and Rating Scales only two are included which stem from the Harvard Clinic.

Research at the Harvard Clinic has aimed at the development of personology, or the clinical study of persons. Most inclusively personology is the study of the life-histories of individual men and of the factors that shape them. It uses the methods of psychoanalysis, individual psychology, and the experimental psychology of personality. Murray does not like the phrase psychology of personality because it is "clumsy and tautological." "Our work is the natural child of the deep, significant, metaphorical, provocative and questionable speculations of psychoanalysis and the precise, systematic, statistical, trivial, and artificial methods of academic personology" (p. 33). Murray also expresses a more impartial view of personology, the blend of all these, when he acknowledges with modesty, humor, and (perhaps unconscious) punning that "personology is still in diapers enjoying random movements" (p. 6).

The general procedure at the Clinic was to examine a group of subjects in a series of experimental sessions and planned conferences. These subjects, 51 men of college age, were most of them Harvard students chosen by the Harvard Employment Office so that (in the rather bi-modal phrase of the book, p. xi) high and low scholarship, arts and sciences, were equally represented. Four groups were studied, each group more intensively than the one before it, as new methods were added and old ones perfected or discarded. The final group of 15 subjects gave, or rather were

paid for, some 36 hours of subject service, which, it is pointed out, is much less than is required in standard psychoanalytic practice.

The unifying purpose behind the procedures was "to organize the findings from all of them into an intelligible portrait of a subject" (p. x). The word portrait deserves stress. Murray's introduction to the one case-history, 87 pages in length, for which space in this volume could be found, asserts that the proof of the personological pudding lies in the case-history and "its worth can be judged positively by its general success in ordering the facts as well as by what it reveals beyond the facts" (p. 606).

Much space would be required to detail the order of procedures in putting a subject through the Clinic. The basic ideas were "to have all experimenters study the same series of individuals with the same concepts actively in mind, and then in assembly—a meeting ('usually lasting five or six hours,' p. 29) being devoted to each case—to report their findings and collaborate in accomplishing a common purpose: the formulation of the personality of every subject" (p. vii). An atmosphere extraordinarily favorable to drawing out confidences and imaginative material prevailed in the Clinic: "we respected our subjects and became fond of them" (p. 13)—and, the reviewer adds, gave them an extremely good time in an atmosphere that must often have seemed to them more like a select social club than a laboratory. There may be a touch of significance in the fact that the subjects are referred to in the report by such extraordinary names as Bulge, Mauve, Oriol, Vulner, and Zora. Certainly the section of the book called "Conversations" which describes the use made of the periods between experiments creates a distinctly bad impression by its preciosity and la-di-da overtones.

At the start of work with every *S*, he was interviewed for 45 min. by the Diagnostic Council in a carefully planned and stereotyped conference. The conference is likened to "the prelude to an opera in that it included parts of themes (tests) which were to be fully presented in subsequent sessions." The purpose of the conference was to provide each of the 5 members of the Diagnostic Council his first opportunity for rating the subject, intuitively, on 40 character traits. A few days later Dr. Murray, as chief experimenter, summarized for the other members of the Council his analysis of the interview. This is in accord with "the principle of weighted judgments" adopted by the Clinic. "In any group of experimenters . . . it is advisable to establish a diagnostic hierarchy. By weighting the opinions of the more competent, one gets the full benefit of *superior* judgments as well as of *many* judgments" (p. 27).

In what followed the subject wrote an autobiography, his early childhood memories were phonographically recorded, his sexual development was probed by a psychiatrist using free-association methods, he rated himself on three 200-item questionaries covering personality variables, and then as *S* he was put through some twenty different test situations or experiments. At the conclusion of all these the examiners met to exchange their findings on a given *S* and then one of them, aided especially by the judgments of the Diagnostic Council, prepared an elaborate life-history with the abstract biography or "psychograph" at the end.

Outstanding among the clinic's inventions are a series of tests devised to lay bare repressed trends. In one of them, the Thematic Apperception Test, which is described to *S* as a test of his creative imagination, he is asked to make up a story

about a picture that is shown him. We are told that all of the *Ss*, without exception, responded projectively, thus betraying their "circumstances, experiences or preoccupations." "For instance, in one of the early experiments six of the eleven college men who took the test said that the youth in one picture was a student; whereas none of the twelve non-college men who acted as subjects described him as such" (p. 533). Projection? Well, to a non-Freudian it may seem that the experimenters are the ones who are most surely doing the projecting; they are either projecting projection or renaming association. Can an association psychologist be denounced if he parsimoniously prefers to see in the six boys out of eleven for which the psychoanalytic concept supposedly "clicked" the simple working of "bonds," of frequency and recency, or redintegration? This instance of bias in interpretation is, of course, a trivial one, but it is cited as one example out of hundreds in this book where the validity of psychoanalytic concepts is inferred from the way a handful of straws blows by in the wind.

Like all tests employed at the clinic the Thematic Apperception Test is valued as a help in constructing for each *S* a psychological portrait that is genetic, dynamic, and psychoanalytic. This test in particular "could be most certainly relied upon to supply the necessary clues for the derivation of the unity-thema" (p. 545). The unity-thema—"the most fundamental and characteristic determinant of a personality"—is conceived to be a compound of needs, partly integrated and partly conflicting, traceable to some experience in early childhood. Such a concept is purely Freudian and congruence with it constitutes a new kind of validation for a test.

The unity-thema of one *S*, Earnst, as elaborated in the psychograph (which is the capstone of his long case history), runs like this: A frail infant, strenuously activated by inanition during the oral period helplessly seeks nourishment from his mother's body; but the father, as he comes to believe, robbed him of the nourishment due him by having sexual connections with the mother, an act which he witnessed, according to the present hypothesis, with lasting effects on his character. Deprivation, repeated in varying forms throughout a neglected childhood, leads him to form the habit of aggressive quest for provision; he struggles on, a sort of lone wolf, immersed in fantasies and anxieties, overcoming one handicap after another, until he becomes a graduate student at Harvard and eventually lands a job. If the reader cares, and dares, he may trace the clues, like those in an intricately embroidered detective story which uncover the plot of the biography and its unity-thema. And he will see how the latter when once grasped, "serves deductively in the reinterpretation of these details (*i.e. S*'s memories, fantasies, attitudes, choice of words, etc.) in the light of the unitary pattern they form." Just as the present reviewer, who is a naïve and inexpert reader of detective stories must always go back, once he gets hold of a plot's denouement, to fit in the details which meant nothing to him when he first read them.

Even a skeptic cannot withhold his admiration for the intuitive artistry with which the method is carried through. If only it were as easy for the skeptic as it is for Murray to be sure that "in the wake of intuition comes investigation directed at crucial problems rather than mere unenlightened fact-collecting" (p. 21). Murray does not doubt this. In effect he says, "Except ye first believe ye shall in no wise enter into the Kingdom" of a worthwhile personology. But the stand-offish extraceptive remains a skeptic. It is his observation that what follows oftener in the wake of intuition is more intuition, and in the wake of belief follows theology. To him a

theology is exactly what a polished-off Freudian system looks like, and no art can help it to masquerade successfully as the fruit of sober inductive methods. To the skeptic, also, the Freudians seem like those theologians who find it easier to believe their creations in bulk than one item at a time, and so they make big and elaborate systems and thus insure that by sheer piling up they eventually become imposing.

Murray states that he knows of no case histories that are more scientific than the clinic's, and he objects to "the current tendency among American psychologists to become suspicious whenever there appears in the writings or speech of a fellow-scientist the slightest trace of aesthetic feeling. A psychologist who believes that he can tell the truth without being 'literary' has only to try writing a case history or biography, and then compare what he has done to a character sketch by any novelist of the first order" (p. 608 ff.).

Who can doubt it? Certainly not the reviewer, who has attempted to teach undergraduates a course in "Biographical Psychology" and has largely failed because he has strong scruples against a wholesale use of the Freudian ideology for its "resolving power" (Murray) and in order to conceptualize a life history and to make dramatic sense of it (Dollard). Why does he feel scruples? Simply because even today, almost fifty years since Freud had his first inspiration, no psychoanalyst seems to feel called upon to make a real effort to *prove* his stuff or to show that he has ever heard of scientific method as it is understood in the physical sciences and as Mill conceived it when he formulated his logic of the Joint-Method. Murray says that "depth psychology flowered because trait psychology failed to 'explain' phenomena, failed to get at the root of things." The reviewer would add that depth psychology resulted when certain psychiatrists recklessly, or out of desperation, rushed in where experimentally trained psychologists were loath to tread!

Psychoanalysis since its beginning has consisted in great part of unverified and prescientific intuition promoted by a group of persons who lack appreciation of what is required to establish the validity of sweeping propositions. Why do its supporters fail to see that instead of asserting that this or that generalized law of development is true what we need to be told is how much or how often it is true, and under what conditions; and on the other hand when it is not true, and under what conditions these exceptions appear, etc., *gradus ad scientiam?* Even now we are waiting for a few at least of the psychoanalysts to detach themselves from the group of believers who maintain faith by becoming completely absorbed in the business of taking in each other's washing and getting all excited over cleaning up first this detailed spot and then that one, who will instead make it their business to set about accumulating the *evidence* in a form that will carry conviction to non-believers.

Murray, unfortunately, can hardly be credited with doing much of the latter. Although in places he roundly attacks the shortcomings and errors of the Freudians (cf. esp. pp. 722-725), he has chosen for himself no such forbidding task as a frontal attempt to bring psychoanalysis into line with science. But Murray does wish to see some of the psychologist's techniques added to the repertory of the psychoanalyst. He is out to redress the prevailing balance between too broad speculation in psychoanalysis and too detailed measurement in the psychology of personality. He has little interest in tests as such, but he does see that tests can do something, and by using a great variety of tests and questionaries and experiments

he brings the manifest measurable traits of personality into the picture as no pure psychoanalyst has done; on the other hand he has directed attention to the latent trends of the person to such a degree that his work is a real marriage, and the first one, between experimental psychology and Freudian or depth psychology.

Different judges will undoubtedly form different opinions of the success of this undertaking. To a sizeable group of psychologists interested in personality it will bring invigorating encouragement. This group will include persons whose hope of attaining a worthwhile psychology of the individual person has been flagging because they have become disillusioned with ordinary mental tests, and who, at the same time, feel no distaste for psychoanalytic theory and wide horizons in general. To another and, we suspect, a larger group of psychologists, including all adherents of rigidly controlled experimental methods and carefully used statistics, it will be a trying book. At the extreme its outlook will seem to them unscientific, its methodology undisciplined, and its immediate effect regressive. Psychoanalysis implemented and eked out with tests will to them still be psychoanalysis, and anathema. Psychology, they are convinced, will develop surely and sanely only if it throws itself into today's definite tasks and does not allow itself to become beguiled by far and enticing vistas.

Not so these authors. With courage and ardor they tolerate and even advocate the "intraceptive" approach to psychological problems. An indispensable glossary of neologisms and redefined terms is added at the end of the book. In it intraception is characterized as "an imaginative, subjective, human outlook." Because they admit the intraceptive attitude into the practice of psychology these authors hold that psychology is too young and too complicated to be as sober and meticulous in gathering data as it has hitherto tried to be. "In its present stage personology seems to call for men who can view things in the broad. . . . A man who has been trained in the exact sciences will find himself somewhat at a loss, if not at a disadvantage. He will find it difficult to fall in with the loose flow of psychologic thought. . . . And so if he continues to hold rigidly to the scientific ideal, to cling to the hope that the results of his researches will approach in accuracy and elegance the formulations of the exact disciplines, he is doomed to failure. He will end his days in the congregation of futile men, of whom the greater number, contractedly withdrawn from critical issues measures trifles with sanctimonious precision. . . . For the present the destiny of personology is best served by giving scope to speculation, perhaps not so much as psychoanalysts allow themselves, but plenty. Hence, in the present volume we have checked self-criticism, ignored various details, winked a little at statistics, and from first to last have never hesitated to offer interpretative hypotheses" (p. 21 f.).

Faced with a credo like this, we can safely surmise that all determined critics of this book will, in contrast, belong to the other camp. They will exemplify Extraception, "a disposition to adhere to the obviously substantial facts. A practical 'down-to-earth' skeptical attitude." Such phrases fit the scientist's habitual description of himself.

Now the men, or more probably man, who coined the terms Intraception and Extraception, and so ingeniously, and in so many connections, develops the contrast between them (pp. 212 ff.), and then says quite simply that in his own psychology he desires to correct for the prevalent anti-intraceptive bias, is smart enough to foresee the gauntlet of criticism which *Explorations in Personality* must run. He is

aware in advance what the blows will be like and from what directions they will rain. Yet he has persevered, undismayed and uncowed by the prospect.

The result is a bold and brilliant piece of work. The present reviewer, believing that American psychologists can perfectly well take care of themselves, has felt a strong temptation to set the book up in an entirely favorable light as a first rate creative achievement on the intraceptive side, and altogether a mine of suggestiveness.

In particular, he is impressed by two arduous analytical enterprises that Murray has held himself to. The first is found in Chapter III, one hundred pages long, on the anatomy, or the variables, of personality. Forty-four variables and their subaspects are distinguished. Thirty-two variables are classified as needs, either manifest or latent, and the remainder as general traits or inner factors. A sub-analysis of each variable, such for example as the need for dominance, or the need for deference, indicates for each its associated desires, effects, stimulating conditions, actions, conflicts, etc. The elaborateness of the conceptual exercise thus carried through goes far beyond anything previously attempted. McDougall's analysis, in comparison, is a product of horse and buggy days. It is no wonder that Murray has needed to introduce various neologisms, some of which grate on our unaccustomed ears. Samples are Infavoidance, Blamavoidance, Press—which is both singular and plural in form and is used to designate any stimulus or evoking situation.

That Murray's dynamic of personality derives from Lewin as well as from McDougall is apparent in Chapter V, over a hundred pages in length, which supplements the analysis of the person into drives or needs by an equally necessary analysis of the environment. This analysis is undertaken in order to be able "to formulate a single man-environment event," in which "only factors that are operating at the moment can rightfully be included" (p. 284). The "press" which operate in childhood are classified under 20 main heads. And this is followed by the distinguishing of 36 types of behavioral reaction. When they had completed this difficult task Murray and his co-workers were equipped with a grammar and a vocabulary equal to any emergency. They devised a system of notation for it, memorized it, and applied it with the utmost facility in the preparation of case studies. No one will do this kind of thing better.

It will hardly do, however, to limit comment to the more praise worthy aspects of the book. Vast amounts of both system and detail in it are not given a straightjacket form of organization partly because it may be premature to give such material that kind of treatment, partly because Murray and his colleagues at the outset rejected the idea of doing things in this way, and partly because they couldn't have done it if they had attempted to. In their experimentation they say they "tried out" almost every kind of mental test, yet apparently the criterion for retention of a test in their scheme was that its results fitted in with their intuitions, which are predominantly Freudian, or at least genetic-dynamic. Whether a test had been standardized or not did not particularly interest them. Their own experiments, without apology, are described with insufficient precision to permit other workers to reproduce them, which is the worst charge of all; often one cannot even understand exactly what was done; procedures were altered, if it seemed best, with each change to a new group of *Ss* so that comparative data were rarely secured; small numbers of *Ss* had to be employed (15 in the group studied most intensively) but no question is raised concerning the effect of limited numbers on the significance

of the correlations reported; very few distributions of scores and P.E. are reported, and reliabilities though sometimes given are treated very cavalierly; simple Chi-square tests of significance were not computed, though the reader can here and there in a moment or two spent figuring them either verify or dismiss important claims; niceties of rating-scale technique which are expected today as standard practice were overlooked; surprisingly little attempt was made to reach objectivity through use of the "matching method," though the writers claim that in the mathematical treatment of results they relied chiefly upon Allport and Vernon's *Measurement of Expressive Movements* (p. 30); no results of standard intelligence or achievement tests or widely used and standardized personality inventories and questionaries are incorporated in the final portraits of the *Ss* (except Allport and Vernon's *Study of Values*); vocational choices are meticulously traced to complexes, here called need-integrates, but no attempt was made to throw a more precise light on interest patterns through Strong's invaluable Vocational Interest Analysis Test; typical of the extraordinary preference in the clinic for its own creations is the fact that, though many tests of creative imagination were used, the widely esteemed and elaborately standardized Rorschach Test was administered to but two-thirds of the final group of 15 *Ss*; childhood memories of college-age *Ss* were tapped but no attempt was made to obtain corroborative checks on their recollections through interviews with parents or other associates; and so on, and on.

Yet at the end of the book, Murray "ventures the opinion" that his program "will eventually be most useful to the experimentalist, the very man who at first glance may find it least appealing" (p. 705). In the last paragraph of the book he makes a plea that other centers where there are "six or more competent examiners" should adopt the same general procedure. But from sheer lack of resources few, if any, groups could do things on such a grand scale as was possible at the Harvard Clinic, with the aid they received from the Rockefeller Foundation. Furthermore, the peculiar inspiration at Harvard will almost certainly not be transferable elsewhere. So our best hope is that on its own side the Harvard Clinic will henceforth become less of an asylum for "epicures of the unconscious" (p. 430), and will strive to discover why it should acquire more respect for what others have been doing; and from the outside that investigators will get busy and throw some of the techniques insufficiently worked out at the clinic into straightjackets and tame some of the hypotheses so airily formulated into manageability. As examples, see the method of confirming a triadic hypothesis concerning hypnotizability which connects Janet and Freud (pp. 486 ff.); the level of aspiration test, which, as reported, is long on interpretation and very short on empirical test; and sex differences in the dramatic production test, with results reported for only 5 college girls.

Fortunately Murray does not mean to claim that any of the Clinic's methods and achievements will be lastingly important in the form in which we are offered them in this book. "Now we have come to the end of our explorations and must . . . write our findings in the sand, since we have nothing to carve in stone, our work being merely a prelude to more substantial studies" (p. 703). Well said! And better still this remark, "'very much at sea' will be our critics' comment." Here, and in rather numerous "asides" of a masochistic-defiant nature scattered through the book, we glimpse deference to what most investigators understand as the real meaning of scientific work—a precise methodology collectively devised and pragmatically but impersonally evaluated.

For it is by the objectivity and collectivity of what they do that psychologists will prove to the world that their science too has grown up and has left behind the days when only men who give free rein to empathy and who are startling and impressive in themselves can achieve something significant. I quote C. P. Snow: "For the point about science is not that it has to be done by archangels, but that it is foolproof, in the 'long run,' for ordinary fallible men. . . . The detached intelligence which is part of the legend and admired from outside with an almost superstitious awe happens to be supplied by the organization of the scientific process and not by the people who perform it."

Had this view of scientific work been kept in mind at the Harvard Clinic the present volume would have been quite different from what it is. A much more modest selection of problems would have been worked on with far more dogged thoroughness. A statistician would not have been called in now and then as consultant with his "instruction" and "guidance" acknowledged in a footnote, but his name not included among the 28 authors on the title page. Instead, the statistician from the beginning of the whole project would have been appointed assistant chief of staff and he would have been armed with extensive veto power. His duties would have included an especially imperative one: to provide rigorous statistical training for the graduate students in psychology who were associated with the project. He would have taught them what Pope, by changing two words, can be made to say.

> "A little *statistic* is a dangerous thing
> Drink deep, or taste not *Personality's* spring,
> There shallow draughts intoxicate the brain
> And drinking largely sobers us again."

It would also have been recognized at the Clinic that a spark of "intuition" can energize months and years of follow-up devoted to forging and validating the instruments of personality analysis; that a single instrument of this kind built into the continuing coöperative tradition is a really great accomplishment.

Explorations in Personality has not been written with these ideals in mind, yet it does something that it is fascinating to see done and does it more colorfully and daringly and with more energy than anyone would have believed possible. We must say quite plainly, however, that one book like this is enough. We wish that there were in it more signs that the promised sequel will be devoted almost 100% to the development of decent technologies dealing with a very small range of the problems of human personality—material of the sort published most naturally in the journals.

But *Explorations in Personality* as it stands should be required *weekday* reading for the psychoanalyst and all psychoanalytically inclined psychiatrists, social case workers and others who are struggling with similar problems and who will profit from the experience of a group that is exploring new vistas with new methods. And it should also be required *Sunday* reading for the psychologists whose field is personality, for they also will find in it ingenious hints of all kinds, as well as the constant vision of that far-off goal, the fathoming of *a* man. The psychologist is faced with a task too overwhelmingly great to justify him in refusing, either on *a priori* grounds, or on the grounds of his brief past experience, to make use of help that is offered him from any quarter, even by an often fantastic-seeming intraceptist. For if he does, then someday Henry Murray may be able to taunt him, or

a Murray disciple in some later day may taunt his imitator, in the words used by Henry IV of France: "Hang yourself, brave Crillon, We fought at Arques and you were not there!"

University of Minnesota RICHARD M. ELLIOTT

NAFE'S VASCULAR THEORY AND THE PREPONDERANCE OF EVIDENCE

In a recent issue of this JOURNAL,[1] Dr. Nafe has replied point-for-point to my criticism of his theory.[2] Instead of a similarly detailed rebuttal, it seems better to re-weigh the evidence pro and con, adding some new points which have developed since the original criticism was written.

According to Nafe, there are no specific thermal receptors. Cold is a kinesthetic pattern from contracting blood vessels; warmth from vasodilation. The chief support for this hypothesis comes from research studies by Nafe and his colleagues. Let us examine these experiments briefly. The fact that thermal sensitivity and blood vessels are both absent from the cornea,[3] is merely argument-by-exclusion—exceptionally weak in this case because the cornea is devoid of so much else besides blood vessels. The apparently coincidental failure of vasoconstriction and cold sensitivity under prolonged stasis is arresting,[4] but it must be remembered that these painful experiments were few in number—too few to establish the exact temporal relation. The plethysmographic studies showing a parallel between vascular activity and thermal experience suggest but do not establish that there is a causal connection.[5] Deaths by drowning and the consumption of ice cream also rise and fall together, but whether either is the cause of the other might reasonably be doubted.

To Nafe's own research may be added Goldscheider's histological study showing the presence of blood vessel clusters under warm and cold spots,[6] although Nafe and I differ as to the meaning of some of Goldscheider's statements and the significance of work performed in 1886 and unconfirmed since. Another good point for the vascular theory is the 'reflex cold' from screeching sounds. It should be noted, however, that Krogh interprets a similar experience from stimulation of the pharynx in terms of a sudden sensitization of specific cold receptors by sympathetic impulses.[7] Nafe would also cite the inadequacy of Von Frey's theory of capsulated thermal receptors (Krause end-bulbs and Ruffini cylinders), which he seems to

[1] J. P. Nafe, Dr. W. L. Jenkins on the vascular theory of warmth and cold, this JOURNAL, 51, 1938, 763-769.

[2] W. L. Jenkins, A critical examination of Nafe's theory of thermal sensitivity, *ibid.*, 51, 1938, 424-429.

[3] J. P. Nafe and K. S. Wagoner, The sensitivity of the cornea of the eye, *J. Psychol.*, 2, 1936, 433-439; The sensitivity of the cornea to heat and pain derived from high temperatures, this JOURNAL, 49, 1937, 631-635.

[4] Nafe and Wagoner, The dependency of cold upon vascular action: studies with nerve block, this JOURNAL, 49, 1937, 636-644.

[5] Nafe and Wagoner, The effect of thermal stimulation upon dilation and constriction of the blood vessels of the contralateral hand, *J. Psychol.*, 2, 1936, 461-477; The effect of adaptation upon vascular reactions to thermal stimuli, this JOURNAL, 49, 1937, 645-649.

[6] Alfred Goldscheider, Histologische Untersuchungen über die Endigungsweise der Hautsinnesnerven beim Menschen, *Arch. f. Anat. u. Physiol.*, Physiol. Abt., Supl. Bd., 5, 1886, 191-227.

[7] A. Krogh, Sympathetic innervation of the cold spots, brought about in a reflex from the pharynx to the skin in man, *Skand. Arch. Physiol.*, 71, 1934, 1-6.

consider is the only alternative to his vascular hypothesis. Surely this is not the case. Other views are possible. In a recent publication, for example, the suggestion is made that the gradations of thermal sensitivity found in seriatim mapping depend upon density patterns of minute receptors and not upon any sort of relatively gross endings corresponding to warm and cold spots.[8]

The supporting evidence for Nafe's theory may now be summarized under two headings: (1) Physiologically, thermal experience parallels vascular activity, positively (Nafe's plethysmographic studies, 'reflex cold') and negatively (Nafe's stasis experiments). (2) Anatomically, thermal sensitivity coincides with the presence of blood vessels, positively (Goldscheider) and negatively (Nafe's corneal studies). In the absence of contrary data, these points would constitute strongly presumptive, though scarcely conclusive, evidence for the vascular theory.

Once we leave this narrow realm, however, doubts begin to emerge. Nafe's theory is essentially one of great simplicity. The general trend of research in thermal sensitivity is in the direction of continually increasing complexity. Let us consider two examples:

(1) Adaptation used to be a very simple affair, explained by Weber as the attainment of physical temperature equilibrium and by Hering as a shift in the physiological zero. Recent work, however, shows that it is far from simple. With large stimulus areas, adaptation is typically slow and incomplete. With very small stimulators, on the other hand, complete adaptation can occur in a few seconds,[9] even with widely spaced stimulations.[10] Rapid restoration of cold sensitivity can be effected by application of a high temperature.[11] Slowly lowering the temperature of the skin surrounding a small area causes the area to lose sensitivity to cold quite suddenly when a certain mild critical temperature is reached; whereas slowly increasing the surrounding temperature to a mild degree sometimes heightens, sometimes depresses sensitivity.[12] Not only the area, but also the form of the stimulator may be important. In some cases, the time for complete adaptation is proportional to the perimeter, in others to the square root of the area, and mixed effects can also be found.[13]

(2) The distribution of sensitivity to warm and cold was formerly a very simple matter. There were so many warm spots and so many cold spots per unit area—supposedly corresponding to individual receptors. Seriatim mapping, however, discloses many gradations of sensitivity—hills and valleys, as it were, instead of isolated warm and cold spots.[14] This sort of distribution is found even when the mapping is

---

[8] Jenkins, Studies in thermal sensitivity: 9. The reliability of seriatim cold-mapping with untrained subjects, *J. Exper. Psychol.*, 24, 1939, 278-293.

[9] S. Aronoff and K. M. Dallenbach, Adaptation of warm spots under continuous and intermittent stimulation, this JOURNAL, 48, 1936, 485-490; H. A. Levine and K. M. Dallenbach, Adaptation of cold spots under continuous and intermittent stimulation, *ibid.*, 48, 1936, 490-497.

[10] Jenkins, Adaptation in isolated cold spots, this JOURNAL, 1937, 49, 1-21.

[11] *Ibid.*

[12] *Ibid.*

[13] Jenkins, Studies in thermal sensitivity: 2. Adaptation with a series of small rectangular stimulators, *J. Exper. Psychol.*, 22, 1938, 84-89.

[14] Jenkins, Studies in thermal sensitivity: 9. The reliability of seriatim cold-mapping with untrained subjects, *J. Exper. Psychol.*, 24, 1939, 278-293.

done with a 1-mm.-diam. stimulator. Alternate mapping with stimulators of different sizes likewise shows complex relationships in the matter of summation effects.[35]

To explain these complicated phenomena in terms of simple vasoconstriction and dilation taxes the imagination. A more logical inference is that the underlying processes are also complex. The vascular theory is not definitely disproved by such facts; it is simply rendered far less credible.

Let us take another broad viewpoint. Nafe's theory postulates a single anatomical mechanism (blood vessels) for the mediation of warmth and cold. Most research points to a dual mechanism. The simple fact that some parts of the skin show only cold sensitivity, others only warm sensitivity, makes it probable that there are separate receptors for the two modalities. Abnormal stimulation (paradoxical warm and cold; excitation by electrical, mechanical and chemical means) also implies separate receptors which are specific in the same manner as those of vision, audition, etc. That is, cold receptors which always give cold, warm receptors which always give warm—no matter how they happen to be stimulated.[36] This is the natural outcome of a dual system, but with the single anatomical unit of the vascular theory, some difficulty is encountered. Either this apparent specificity must be proved fallacious, or else blood vessels must mimic the action of specific receptors. That is, some of them must always constrict and others always relax, regardless of the kind of stimulation. Nafe cites evidence to show that reversals of normal vascular action do sometimes occur (dilation when we should expect constriction and vice versa), but this is scarcely proof that blood vessels act consistently in the specific manner called for.

Also hard to explain in vascular terms is Lanier's finding that warm sensitivity is regenerated much more slowly than cold, following nerve block by alcohol injection.[37] Such dissociation seems incredible with an anatomically-single system. In terms of Nafe's hypothesis, it would have to mean that the blood vessels in the affected area had recovered the ability to constrict, but not to relax—or else that the impulses from constriction are carried by different nerve fibers from those of dilation. Neither view is particularly plausible.

Finally, the factor of *time* must be briefly considered. We may agree with Nafe that the matter of reaction time should be discarded, because of uncertainty as to some of the most critical components.[38] There remains, however, the problem of the

---

[35] Unpublished studies by the author.

[36] Nafe uses the term 'specific' to mean that a receptor can *receive* only one kind of stimulation. Thus he argues that paradoxical experiences, electrical stimulation, etc., show that thermal receptors are *not* specific. This misunderstanding is unfortunate. No one would undertake to maintain that thermal receptors are specific in Nafe's sense. The whole point is that they can be stimulated in a number of ways—but always with the same end result. Actually, these abnormal forms of stimulation are the best evidence that the experience depends upon the receptor stimulated and not upon the nature of the stimulus.

[37] L. H. Lanier, An experimental study of cutaneous innervation, *Proc. Ass. Res. Nerv. & Ment. Dis.*, 15, 1935, 437-456.

[38] Reaction-time still remains a potentially lethal enemy for the vascular theory. Obviously, if there is not time enough for the necessary events to occur, the theory must be false, no matter how plausible the arguments in its favor. Nafe's own computed figure of 250 ms. for the minimum in terms of vascular action is still too large to check with some of the reaction times to cold reported in the literature.

brief critical flicker frequency for intermittent warm stimulations. Geblewicz[10] has recently confirmed Kastorf's figures by more exact methods. Successive warm stimulations, he shows, separated by only 0.15 sec. can be sensed as discrete. If intermittent radiant stimulation caused *alternating* warmth and cold experiences, what support this would provide for Nafe's hypothesis! Unfortunately, such is not the case. Only flickering *warmth* is experienced. To interpret this rationally in vascular terms seems almost impossible. Not only must a cycle of relaxation and constriction occur within 0.15 sec. (unless we assume that blood vessels can keep on relaxing without intervening constrictions)—but the *contraction phase must occur too slowly to be sensed as cold!*

If we now step back and view the evidence as a whole, the overwhelming weight on the negative side becomes apparent. Nafe may justly maintain that no one of these facts absolutely proves his theory to be wrong. That is true. It is not the individual difficulties but the everlasting succession of them that is damning. To justify its acceptance, a theory should have the majority of experimental findings in its favor. In this case, the *preponderance* of evidence is strongly against the vascular hypothesis.

Lehigh University WILLIAM LEROY JENKINS

THRESHOLDS OF VISUAL INTENSITY-DISCRIMINATION IN PHYLOGENY

A recent paper by Ghiselli established the point that the principle of encephalization of function does not appear to hold in phylogeny regarding intensity discrimination in vision.[1] It was pointed out that in the mammalian series the cortical centers for brightness discrimination were of as great functional importance in lower (rat[2]) as in higher (cat,[3] dog,[4] monkey[5]) forms. In all mammals, except man, the subcortical visual centers seem able to mediate the brightness habit to some extent in the absence of the striate area of the cortex. Since there seems to be little change in the functional importance of the retinal projection centers in the mammalian series, although anatomically there is a gradual telencephalization of structure,[6] it is of interest to examine the behavioral data regarding the discriminative capacity of different vertebrate organisms in an effort to determine whether this structural modification enables the animal to differentiate brightnesses more keenly.

---

[10] E. Geblewicz, L'influence de l'intensité sur las persistance apparente des excitations thermiques pour deux stimulations successives, *C.r. Soc. Biol.*, 127, 1938, 173-175.

[1] Edwin Ghiselli, Encephalization of brightness discrimination in mammals, *Science*, 86, 1937, 618-619.

[2] K. S. Lashley, The mechanism of vision: II. The influence of cerebral lesions upon the threshold of discrimination of brightness, *Ped. Sem.*, 37, 1930, 461-480.

[3] K. U. Smith, Visual discrimination in the cat: V. The postoperative effects of removal of the striate areas upon intensity discrimination, *ibid.*, 51, 1937, 329-370.

[4] D. G. Marquis, Effects of removal of the visual cortex in mammals, with observations on the retention of light discrimination in dogs, *Proc. Ass. Res. Ner. Men. Dis.*, 13, 1934, 558-592.

[5] Heinrich Klüver, An analysis of the effects of removal of the occipital lobes in monkeys, *J. Psychol.*, 2, 1936, 49-61; D. G. Marquis, Phylogenetic interpretation of the functions of the visual cortex, *Arch. Neurol. Psychiat.*, 33, 1935, 807-815.

[6] Marquis, *op. cit.*; C. U. Ariëns Kappers, G. C. Huber, and E. C. Crosby, *The Comparative Anatomy of the Nervous System of Vertebrates*, 2, 1936, 1061-1205; J. F. Fulton, *Physiology of the Nervous System*, 1938, 347-365.

Differential brightness limens have been determined for a number of different organisms using the Yerkes-Watson apparatus,[7] or some modification of this experimental arrangement. This is fortunate for the present discussion because the visual intensity thresholds may be directly compared. In the accompanying table all thresholds reported for the infra-human forms, except Frolov's for the dog,[8] were obtained with roughly comparable experimental set-ups. The difference threshold has been expressed in a number of different ways in the literature, and the writer has thought it expedient to convert all thresholds to a common expression, namely, the Weber ratio. An asterisk (*) in the $\Delta I/I$ column indicates that the writer has calculated $\Delta I/I$ from the data given by the investigator. The last column of the table is the brightness level at which the ratio was found. In those cases where the threshold was found at a number of different intensity levels (Crawford,[9] Mead,[10] Moody,[11] Munn-Slater[12]) the intensity at the minimum limen is given. In the experiments by Lashley[13] and Ghiselli,[14] wherein a large number of animals were tested, the minimum threshold found is presented.

Beginning with the mammals, and disregarding the results of Yerkes' experiment with the dancing mouse in which a somewhat easy criterion of discrimination was used,[15] it will be seen that there is a gradual decrease in the threshold value as the phyletic scale is ascended. Among mammals the ability to discriminate brightness differences increases according to the phylogenetic position of the animal in the series. Among the studies included, the mouse and the rat have the poorest vision. These organisms are exceeded in capacity by the raccoon, cat, and dog, whereas the discriminative ability of the monkey lies between this group and that of the human individual, the latter exceeding by far all other animals. The extremely low threshold reported by Frolov on the dog was obtained by a conditioning method and with reflected light.[16] The other infra-human experiments in the table made use of a discrimination method and transmitted light. The superiority of the human is undoubtedly partly a function of the bipartite test field with which the thresholds were determined, these conditions presenting an optimal state for brightness discrimination.[17] In those cases where the human threshold was measured under the same conditions as that of the animals, although the human subject remained superior, the superiority was not as marked.

We may conclude, from an examination of Table I, that there seems to be an

---

[7] R. M. Yerkes and J. B. Watson, Methods of studying vision in animals, *Behav. Monog.*, 1, 1911, No. 2, pp. 90.

[8] G. P. Frolov, Reaction of the central nervous system to changes in intensity of illumination, *Proc. Petrograd Soc. Nat. Sci.*, 69, 1918, 3-19.

[9] M. P. Crawford, Brightness discrimination in the Rhesus monkey, *Genet. Psychol. Monog.*, 17, 1935, 72-162.

[10] L. C. Mead, The curve of visual intensity discrimination in the cat, *Psychol. Bull.*, 34, 1937, 517; and other results as yet unpublished.

[11] P. A. Moody, Brightness discrimination in the deer-mouse—Peromyscus maniculatus gracilis, *J. Exper. Zoöl.*, 52, 1929, 367-405.

[12] N. L. Munn and J. E. Slater, A note on brightness vision in the white rat, *J. Comp. Psychol.*, 13, 1932, 273-277.

[13] Lashley, *op. cit.*

[14] E. E. Ghiselli, The superior colliculus in vision, *J. Comp. Neurol.*, 67, 1937, 451-467.

[15] R. M. Yerkes, *The Dancing Mouse*, 1907, 91-132.

[16] Frolov, *op. cit.*

[17] J. R. Smith, Spatial and binocular effects in human intensity discrimination, *J. Gen. Psychol.*, 14, 1936, 318-245.

increase in the capacity for brightness discrimination from lower to higher mammalian forms. Thus, although the extirpation studies indicate that there is no gradual

TABLE I

COMPARATIVE SUMMARY OF BRIGHTNESS DISCRIMINATION THRESHOLDS

Subject	Investigator†	Minimum ΔI/I	Intensity
Sunfish	Reeves	3.0*	?
Sparrow	Tugman	.18*	.098 cp.
Mouse	Yerkes	.11*	4.1, 16.6, 66.4 cp.
Deer-mouse	Moody	.83*	36.11 cp.
Rat (hooded)	Lashley	1.0*	?
Rat (white)	Munn-Slater	.59*	3.2 ml.
Rat (white)	Slater	.47*	?
Rat (white)	Spencer	.43*	?
Rat (both)	Ghiselli	1.00*	?
Raccoon	Munn	.27*	13.45 ml.
Cat	Smith	.23*	.49, 50.7 ml.
Cat	Mead	.19	.126, 26.1 ml.
Dog	Frolov	.02	?
Dog	Stone	.16*	1 cp.
Dog	Marquis	.20*	?
Monkey	Crawford	.098	.77 ml.
Human	König-Brodhun	.02	10–100 ml.
Human	Steinhardt	.01	1000 ml.
Human	Craik	.01	4000 equiv. fc.

† The references from which the data for this table were obtained are: C. D. Reeves, Discrimination of light of different wave lengths by fish, *Behav. Monog.*, 4, 1919, 1–106; E. F. Tugman, Light discrimination in the English sparrow, *J. Anim. Behav.*, 4, 1914, 55–109; Yerkes, *op. cit.*; Moody, *op. cit.*; Lashley, *op. cit.*; Munn and Slater, *op. cit.*; J. E. Slater, Brightness vision in the albino rat, *Univ. Pitt. Bull.*, 7, 1931, 394–395; L. T. Spencer, Central inhibition in the albino rat, *J. comp. Psychol.*, 3, 1923, 389–408; Ghiselli, *op. cit.* (footnote 14 above); N. L. Munn, Pattern and brightness discrimination in raccoons, *Ped. Sem.*, 37, 1930, 3–34; K. U. Smith, *op. cit.*; Mead, *op. cit.*; Frolov, *op. cit.*; C. P. Stone, Notes on light discrimination in the dog, *J. Comp. Psychol.*, 1, 1921, 413–431; Marquis, *op. cit.*, (footnote 4 above); Crawford, *op. cit.*; A. König und E. Brodhun, Experimentelle Untersuchungen über die psychophysische Fundamentalformel in Bezug auf den Gesichtsinn, *Zitzber. d. Akad. d. Wiss.*, 1889, 641; J. Steinhardt, Intensity discrimination in the human eye: I. The relation of ΔI–I to intensity, *J. Gen. Physiol.*, 20, 1936, 185–209; K. J. W. Craik, The effect of adaptation on differential brightness discrimination, *J. Physiol.*, 92, 1938, 406–421.

shift in function from subcortical to cortical visual centers for this visual capacity in mammals, the anatomical encephalization of structure is accompanied by increased brightness discrimination capacity. The writer suggests that it may be the changes in the retina rather than the neurological modifications which determine this increased visual capacity in phylogeny.

University of Rochester LEONARD C. MEAD

A DEFINITION OF BOREDOM: A REPLY TO MR. BERMAN

In the April issue of this JOURNAL,[1] Berman distorted a definition of boredom which I had suggested,[2] and proceeded to challenge my views within the framework of this distortion. I defined *boredom:*

[1] A. Berman, The effect of benzedrine sulfate (amphetamine sulfate) on satiation, this JOURNAL, 52, 1939, 297-299.

[2] J. E. Barmack, Boredom and other factors in the physiology of mental effort, *Arch. Psychol.*, 31, 1937, (no. 218), 1-83; The effect of benzedrine sulfate (benzyl methyl carbinamine) upon the report of boredom and other factors, *J. Psychol.*, 5, 1938, 125-132.

"as a state of conflict between the tendency to continue and the tendency to get away from a situation which has become unpleasant principally because of inadequate motivation resulting in inadequate physiological adjustments to it. Accordingly, a state of boredom is initiated by inadequate motivation during the operation of a task set and results in a tendency for the physiology of the subject to revert back to the sleep level. The inadequate vital adjustments to the task are unpleasantly appreciated as the feeling of monotony or fatigue. If the task set is weak, the subject may go off to sleep or abandon the task. If the task set is sufficiently strong, the subject struggles to remain awake or partly escape from the depressing task. These latter objectives are achieved usually unconsciously, by shifting attention away from the task, daydreaming, creating extrinsic goals, modifying the procedure, etc."[3]

This view of boredom is elaborated in an early publication.[4] Berman, without explicitly stating that he is using his own terminology, substitutes the word *satiation*[5] for *boredom* in his article which opens as follows:

"In a recent paper *satiation* was defined by Barmack as 'a state of conflict between the tendency to continue and the tendency to get away from a situation.' "[6] (Italics mine.)

There are at least two objections to this type of substitution. First, although the substitution may be quite harmless in one context, it can be unfortunate in another. For example, it is embarrassing to find attributed to me the view which Berman's following statements suggest.

"This is certainly true in certain forms of satiation which we consider primitive; where the satisfaction of a basic need is involved, and is well illustrated by the work of Katz on hunger and appetite. Barmack adopts just this point, *but errs when he generalizes this restricted view to all satiation.*"[7] (Italics mine.)

Secondly, the substitution of the word 'satiation' for 'boredom' is contra-indicated by the procedure of the experiments to which he refers. In belated correspondence, Berman attempted to justify the substitution on the grounds that Lewin and Karsten use the term, 'psychic satiation,' in describing similar experiments.[8] A basic difference in procedure between the experiments of Lewin and Karsten and my own is that in the former the subject is relatively free to stop an activity against which he has developed a negative valence,[9] and in the latter the subject is acting under the constraint of the experimenter's instructions to complete a specified period of activity. In the Lewin and Karsten experiments, the rejection of the activity may be considered a manifestation of satiation. In my own studies and other studies in which boredom has been the focal point of interest, the initial activity may be described conceivably as satiating, but the subsequent activity may be more effectively described as a not too happy *toleration* of an unpleasant task. In this situation, *S* conforms outwardly to the instructions of the *E*, but expresses a negative valence to parts of, or to the whole of the task by a diminished or

---

[3] *Ibid.*, 125 f.

[4] *Op. cit., Arch. Psychol.*, 67-74.

[5] I do not use the word "satiation" in either of the publications to which Berman refers.

[6] *Op. cit.*, 297.

[7] *Ibid.*

[8] K. Lewin and A. Karsten, Untersuchungen zur Handlungs- und Affektpsychologie: V. Psychische Sättigung, *Psychol. Forsch.*, 10, 1928, 142-154.

[9] Lewin describes the technique as follows: "the subject must do a certain task repeatedly; he is, however, free to stop as soon as he has enough of it." *A Dynamic Theory of Personality*, 1935, 254.

variable rate of work, daydreaming, alterations in procedure, etc. 'Psychic satiation,' therefore, may be regarded as but one aspect of the state of boredom.

Again, the substitution of the word (and concept of) 'satiation' has led Berman into what may be regarded as one of the more serious misstatements of my views. He writes:

"When it is observed that an individual is highly motivated, then with continued activity becomes disinterested, and finally rejects the activity entirely, it is plausible to consider that the motivation has somehow disappeared; that in some way the tissue needs, states of agitation, or other conditions which are the basis of motivation (drive) have become appeased, and then saturated to the point of aversion. . . . Barmack adopts just this point, but errs when he generalizes this restricted view to all satiation."[10]

In the quotation from my Benzedrine study appearing in this article, and to a more detailed extent in my other study,[11] it is clear that I refer to the physiological *consequences* of certain motives rather than to their antecedents. These *results* of motivation favor an alert attitude during the work. According to our hypothesis, then, after the exploratory drive is gratified (of which the relative novelty of the task may be considered the incentive), and under certain other conditions such as the lack of the operation of other strong motives, and with the restriction of attention to the task,[12] a condition develops which is unfavorable to the maintenance of the alert state. Out of this tendency stem two main effects, a desire to get away from the task or other environmental condition which, in a sense, had produced it, or a desire (usually unconscious) to correct the unpleasant state by introducing new methods of work, thinking of more strongly affective situations, which are, in a way, means of staying alert, etc. Through conditioning, the range of stimuli to which the rejection may appear is extended.

On the basis of the foregoing analysis, it follows that certain antihypnotics should allay or minimize the appearance of these two effects.[13] In fact, the antihypnotic, Benzedrine, was found effective in retarding (a) a drop in the rate of addition over a two-hour period, (b) the development of boredome, strain, irritation, fatigue, sleepiness and inattentiveness, as indicated by ratings on ten-point subjective rating scales.[14] Subsequently, we obtained evidence that 60 mg. of ephedrine hydrochloride and 2 gr. of caffeine (alkaloid form)[15] have similar properties, although, under the conditions of the experiments, to a lesser degree than the 10 mg. of Benzedrine.

Berman reveals still another confusion from his substitution of the word 'satiation' for 'boredom' as is apparent in the following:

"As is well known, the requisite for sleep, or a sleep-like state, is the exclusion of sensory stimulation. An abundance of stimuli in the field, as would be required to satiate an individual, would tend to reduce the tendency to go to sleep."[16]

It is not true that the requisite for sleep is the exclusion of sensory stimulation. So far as I know, no study has been performed in which the relative importance of motivation or other internal conditions has been compared to that of the exclusion of sensory stimuli, and to describe the conditions which facilitate the development

---

[10] *Op. cit.*, 297. [11] *Op. cit., Arch. Psychol.*, 67-74. [12] *Ibid.*, 71.

[13] Those in which there are no other intrinsic factors which might induce a negative valence to the task.

[14] *Op. cit., J. Psychol.*, 5, 1938, 125-133.

[15] Unpublished material. [16] *Op. cit.*, 298.

of boredom as "an abundance of stimuli in the field" is not altogether warranted. Towards the end of his paper Berman states:

"This note is in part directed to just this recognition of the confusion between the motivational and physiological aspects of satiation on the one hand, and the psychical on the other. It does not posit a strict dualism between motivation and satiation (for this is obviously incorrect), but attempts to indicate *that all aspects of satiation cannot be explained on the grounds of basic motivational states, and tendencies to revert to a sleep level.*"[17] (Italics mine).

Berman, I was informed by correspondence, bases the italicized statement on the following quotation appearing at the end of my Benzedrine study. "A state of boredom is caused principally by a physiological reversion to the sleep level due to inadequate motivation during the operation of a task set."[18]

Berman criticized the use of the subjective rating scales. "We doubt," he writes, "whether any *S* has the ability to discriminate accurately over a 10-point scale. It was found in one of our experiments that responses to a 5-point scale were not sufficiently reliable to be the sole criterion of satiation."[19]

I have not used the subjective rating data as the *sole* criterion of boredom, but have used them in conjunction with changes in rate of addition. In the light of correspondence with Berman, it appears that his statement on the 'unreliability' of the 5-point rating scale which he used was based on an observed inconsistency between certain verbal reports of some individual *Ss* and their ratings. That the ratings of some individuals may not be too valid an expression of their subjective states is common knowledge. In the Benzedrine study, I dealt with *average ratings* of *many Ss* rather than with a consideration of individual cases. It is true that the critical ratios of differences between the means of the subjective ratings for Benzedrine and placebo at specific periods are only of the order 0.95 to 2.5. The duplication, however, of these values for a series of four sets of differences during the last hour of the 2-hr. work-period for each of the dependent variables gives most of them added statistical significance. Certainly the lowest qualitative evaluation that may describe such data is that they cannot be ignored. Further, essentially the same findings were obtained in another experiment using 15 mg. of Benzedrine on 10 *Ss* engaged in an entirely different task, operating the Poffenberger pursuitmeter.[20]

In a better founded criticism, Berman inquires why the data of the Benzedrine and other experiments using sympatho-mimetic compounds should be used to support an hypothesis on boredom rather than a more obvious fatigue interpretation. A feeling of fatigue is, in fact, usually associated with boring work. Because of its transient character, however, it is generally referred to as a pseudo-fatigue or feeling of monotony.

For the following reasons, the data may be more accurately integrated about a boredom concept than about a fatigue concept: (1) The *Ss* themselves verbalized post-experimentally their experiences as predominantly boring, rather than fatiguing, and this was particularly true after the 2-hr. pursuit-task. (2) In another study, *Ss* added numbers under the same conditions for 4 hr. without appreciable difficulty.[21]

---

[17] *Op. cit.*, 299. [18] *Op. cit., J. Psychol.*, 132. [19] *Op. cit.*, 298.
[20] *J. Exper. Psychol.*, (in press).
[21] Barmack, The length of the work-period and the work-curve, *J. Exper. Psychol.*, 25, 1939, 109-115.

It is hardly likely, therefore, that the *Ss* in the Benzedrine experiment "were probably forced to exhaustion" within 2 hr., as Berman suggests.[22] (3) In still another study, with pursuit activity as the task, the introduction of a small added financial incentive after one hour of work resulted in a marked improvement in accuracy of performance for the second hour which surpassed that of any period of the first hour.[23] Accompanying the marked improvement in pursuit performance were equally striking changes in report of boredom, fatique, sleepiness, etc., in the more favorable direction. It is improbable that such changes would occur as readily in fatigued *Ss* as in bored *Ss*. This finding is relevant here because the observed effects of Benzedrine on the report of boredom and other factors were duplicated with this type of task.

The inference is compelling that Berman's article arose chiefly out of a number of confusions which resulted from his unwarranted substitution of both the word and concept of 'satiation' for 'boredom.'

College of the City of New York JOSEPH E. BARMACK

SATIATION OR BOREDOM? A REJOINDER TO DR. BARMACK

In a former paper appearing in this JOURNAL,[1] I attempted to demonstrate that the concept of satiation cannot entirely be explained in terms of a reduction in motivation and return to a sleep level, as suggested by Barmack.[2] The arguments presented are attacked by Barmack.[3] The principal criticism leveled against me is that I accepted the equivalence of the concepts of satiation and boredom—even to the point of a substitution of terms. Before the validity of such a procedure can be determined the operations governing the two concepts must be investigated. The term *psychical satiation* was originally used (as far as it is possible to tell) by Karsten and Lewin.[4] I objected to the term *psychical* because it implies that satiation is purely a mental process. It is easily shown that such a distinction cannot safely be made, for satiation is a behavioral phenomenon and not solely mental. Consequently the appendage *psychical* was dropped and the term *satiation* itself was employed as denoting behavioral satiation. It was recognized, in addition, that satiation has as many manifestations as there are behavioral activities. Thus one can be satiated with eating a certain food, or with all food (*i.e.* hunger satiation); with sexual activity; with listening to music; with adding numbers; with adjusting a bolt on a moving assembly line; etc. In fact, satiation may appear for any task or activity when it is repeated sufficiently to induce a negative valence and rejection of the task. It is not known whether the satiation which arises with the continuous activity of inserting pegs into a board is the same as that which develops with the repetitive ingestion of a single food, or class of foods. It is reasonably certain, however, that in general the principle of the satiation of behavior holds for all levels of activity. In my paper, the satiation of basic drives, *i.e.* hunger, sex, etc., was called primitive as

---

[22] *Op. cit.*, 299. [23] Unpublished material.

[1] A. Berman, The effect of benzedrine sulphate (amphetamine sulphate) on satiation, this JOURNAL, 52, 1939, 297-299.

[2] J. E. Barmack, The effect of benzedrine sulphate (benzyl methyl carbinamine) upon the report of boredom and other factors, *J. Psychol.*, 5, 1938, 125-133.

[3] *Supra*, 467-471.

[4] K. Lewin and A. Karsten, Untersuchungen zur Handlungs- und Affektpsychologie: V., Psychische Sättigung, *Psychol. Forsch.*, 10, 1928, 142-254.

compared to the behavior manifested in the higher levels. In the studies under discussion satiation is confined to behavior above that of the simple drives.

The tasks employed by Dr. Barmack (adding 6-place numbers) and by me (maze activity, peg insertion, and drawing faces)[5] are not generically different from those used by Karsten and Lewin,[6] Curran and Schilder,[7] Fritsche,[8] and Reichenberg[9] in their work on satiation. Yet Dr. Barmack maintains that his experiments deal with boredom and not with satiation.[10] He bases his claim upon the point that his *Ss* were given an arbitrary time limit of 2 hr. to perform; whereas in my experiments the *Ss* were permitted to stop of their own accord; and in later experiments forced to continue until the task was rejected two or three times. He states:

> In my own studies and other studies which boredom has been the focal point of interest, the initial activity may be described conceivably as satiating, but the subsequent activity may be more effectively described as a not too happy *toleration* (italics his) of an unpleasant task. In this situation, the subject conforms outwardly to the instructions of the experimenter, but expresses a negative valence to parts of, or to the whole of the task by a diminished or variable rate of work, day-dreaming, alterations in procedure, etc. Psychic satiation, therefore may be regarded as but one aspect of the state of boredom.[11]

It seems that the entire distinction involves one of degree. When an *S* persists through an arbitrary experimental requirement beyond the point where he would normally reject the task, Dr. Barmack would have us call it boredom; but if *S* rejects the task of his own accord before the time limit, Dr. Barmack would call it satiation. Certainly the task becomes more intolerable as satiation progresses, and *S* musters all his energy and ingenuity to persist, but does this create a new process, or is it merely the same phenomenon carried to an unusual degree, *i.e.* over-satiation? It seems that Dr. Barmack has introduced an extraneous factor. Surely he has not demonstrated that *satiation* is *"but one aspect of the state of boredom"* (italics mine). In the final analysis, an examination of the operational aspects of satiation and boredom reveal that both processes may be elicited in the same manner and show similar reaction patterns within similar frames of reference.

If it is accepted that satiation and boredom are similar processes, then there is no rigid incompatibility in the actual operational facts of Dr. Barmack's experiments and mine. The difficulty arises in the deductions drawn from the facts and the bias of the interpreter. Not a small part of the difficulty seems to be the level of description or explanation preferred for the interpretation. Since our knowledge of the physiological basis of satiation is scant, I prefer a more psychological description. This does not mean, however, that physiological facts are to be ignored when they are known. On the contrary, motivation (drive) does seem to play a part in satiation but it does not explain all the facts as the theory of a return to

---

[5] A. Berman, The relation of time estimation to satiation (*in press*). Successive reproductions of visually perceived forms, (*in press*). The influence of social factors on the persistence of satiation in pre-school children (*in preparation*).

[6] *Op. cit.*

[7] F. J. Curran and P. Schilder, Experiments in repetition and recall, *Ped. Sem.*, 51, 1937, 163-187.

[8] L. Fritsche, Über psychische Sattigung, *Zsch. Psychiat. Neurol.*, 97, 1937, 291.

[9] W. Reichenberg, An experimental investigation on the effect of gratification upon effort and orientation to reality, *Amer. J. Orthopsychiat.*, 9, 1939, 186-203.

[10] *Supra*, 467-471.

[11] *Op. cit.*

a sleep level would imply. There is no disputation of the action of benzedrine sulphate on satiation, but only in the putative justification that this action has for Dr. Barmack's theory.

It is beyond the limits of this note to respond at great length to the variegated ramifications of the problem under discussion. I can, however, delineate some of the arguments under the following rubrics: (1) Some scattered evidence and personal experience indicates—as mentioned in my previous paper—that sleep, or a a sleep-like state, is best fostered by an exclusion of heterogeneous sensory stimuli.[22] In the distal stages of satiation, the mechanisms (*substitutions*) created by *S* to persist at the task provide the heterogeneous stimulation. Not one of the 122 *Ss* employed in a variety of experiments on satiation has ever succumbed to sleep during the trials, or immediately after them. (2) The use of statistical measures, *e.g.* the mean as used by Dr. Barmack, does not invalidate the argument that subjective rating scales cannot be the sole criterion of the presence of satiation. Such a procedure must always be accompanied by non-resumption of activity, characteristic variability in performance, the presence of *substitutions*, as day-dreaming. (3) Forcing the *Ss* to unusual degrees of over-satiation complicates the picture by increasing the fatigue factor, which should be kept at a minimum in satiation studies. In addition, such abnormal periods of activity introduce extraneous quasi-needs and ignore the individual differences of the *Ss* in the capacity for satiation.

It is gratifying that satiation is assuming more and more a place in the experimental psychological literature. Particularly is this true in light of the fact that satiation has been so surprisingly neglected in favor of the concept of aspiration level. When additional facts become available, it will be possible to speak with greater emphasis about satiation.

University of California ARTHUR BERMAN

THE TENTH ANNUAL SPRING MEETING OF THE EASTERN PSYCHOLOGICAL ASSOCIATION

The tenth annual spring meeting of the Eastern Psychological Association was held at Bryn Mawr College, Bryn Mawr, Pennsylvania, on Friday and Saturday, March 31 and April 1, 1939. There were 660 registered for the meeting which is the largest attendance in the Association's history. Of these, 351 were members and associates of the American Psychological Association, and therefore members of the Eastern Association. A total of 309 guests registered. Bryn Mawr College opened the dormitories and dining halls for the use of the Association from Thursday evening until Sunday morning. As all meetings were held on the campus, the facilities were most excellent for so large a group.

A total of 113 papers were presented. On Friday morning the following sessions were held: Educational Psychology, Sensory Psychology, Brain Function, Attitudes, Applied Psychology, and Experimentally Induced Conflict. A special luncheon, sponsored by the A.P.A. Committee on Displaced Foreign Psychologists, was held on Friday noon. Gordon W. Allport, President of the A.P.A., presided. In the afternoon, sessions were held on the following topics: Abnormal Psychology, Vision and Audition, Animal Motivation, Tests and Measurements, General and Experimental, and Genetic Psychology.

[22] David Katz, *Animals and Men*, 1937, 34.

Round tables were held late Friday afternoon. The group discussing the "Functions of the School Psychologist" was led by T. E. Newland of the Pennsylvania Department of Public Instruction. W. S. Hunter of Brown University was chairman of the round table on "The Law of Effect and the Laws of Conditioning." The Society for the Psychological Study of Social Issues sponsored a discussion on "Critical Problems in the Formation and Change of Attitudes." Daniel Katz of Princeton University was the chairman of the group. The round table on "Mechanisms in Visual Perception—Interrelations of Constancy, Contrast, Conversion and Adaptation" was led by Harry Helson of Bryn Mawr College.

The President of the Eastern Psychological Association, Karl M. Dallenbach, Cornell University, addressed the group in Goodhart Hall on Friday evening. His topic was "Pain: History and Present Status." Following this meeting Bryn Mawr College held a reception for members and guests.

On Saturday morning, papers were presented in the fields of Physiological Psychology, Conditioned Response, Personality, Human Learning, Intelligence, and the Rorschach Tests. At noon a special luncheon was held by the Psychologists' Committee for Aid to Spanish Democracy.

Saturday afternoon was taken up with the annual business meeting and a general session. The President presided at both. D. G. Marquis, Yale University, read a paper on "Recent Advances in the Conditioned Response." O. A. Oeser from the University of St. Andrews, Scotland, presented a paper on "Field Work in Social Psychology—The Study of Unemployment in a Scottish Industrial City." George H. Gallup of the American Institute of Public Opinion addressed the group on "Some Problems in the Measurement of Public Opinion."

At the business meeting of the Association, the following recommendations of the Board of Directors were approved: (1) to hold, in the future, annual meetings two days in length; (2) to authorize the appointment of a committee to study the question of membership and of the organization of the Association (the President appointed the following: D. H. Fryer, Chairman, H. S. Langfeld, from among the general membership, and Theodora M. Abel, Harry Helson, and H. S. Oberly from among the members of the Board of Directors); (3) to empower, since an invitation for the 1940 meeting had not been received, the retiring president, the retiring secretary, and the incoming president to select a time and a place for the next annual meeting; and (4) to contribute $50 to the American Psychological Association's Committee on Displaced Foreign Psychologists.

The Association elected F. L. Wells, Harvard University, to serve as President for 1930-1940; Harry Helson, Bryn Mawr College, to fill the unexpired term of one year as Secretary-Treasurer; Theodora M. Abel, Trade Extension Classes, New York City, and George W. Hartman, Teachers College, Columbia University, to serve as members of the Board of Directors for the period 1939-1942; H. S. Oberly, University of Pennsylvania, to serve as a member of the Board of Directors, during the one year remaining in the term of Harry Helson.

The Board of Directors announced the appointment of the following Committees for 1939-1940. Program Committee: E. Heidbreder, Wellesley College, Chairman, Ross A. McFarland, Harvard University, and L. W. Max, New York University; Nominating Committee: George Kreezer, Cornell University, Chairman, M. B. McGraw, Columbia Medical Center, O. H. Mowrer, Yale University.

The eleventh annual meeting of the association will be held, as was subsequently decided by the Committee appointed for that purpose, at Atlantic City, New Jersey, on Friday and Saturday, April 5 and 6, 1940.

University of Pennsylvania H. S. OBERLY

THE FOURTEENTH ANNUAL MEETING OF THE MIDWESTERN PSYCHOLOGICAL ASSOCIATION

The Midwestern Psychological Association held its fourteenth annual meeting at the University of Nebraska, Lincoln, Nebraska, Friday and Saturday, May 5 and 6, 1939. The register was signed by 362 members and visitors who attended the meeting. On this occasion the University of Nebraska commemorated the 50th anniversary of the founding of its psychological laboratory by Harry Kirke Wolfe.

At the annual business meeting on Friday afternoon, 54 new members were elected, bringing the present membership list to 469. Announcement was made of the election of J. P. Guilford to the presidency for the year 1939-40; of Robert H. Seashore to the office of secretary-treasurer for the term of 1939-42; and of Norman R. F. Maier to the executive council for the term 1939-42. It was voted to contribute $25 to the American Psychological Association's Committee on Displaced Foreign Psychologists. It was announced that the next annual meeting of the Association will be held at the University of Chicago, the exact dates to be announced later.

Ten sessions were held at which 65 scheduled papers were read. The emphasis, as indicated by the numbers of papers, was upon studies of attitudes which occupied two sessions. Other sessions were on the conditioned response, mental development, personality, sensory and nervous phenomena, applied psychology, learning, and mental tests. Symposia were conducted on clinical psychology, personality, mental hygiene, motivation, constancy of the IQ, contemporary social problems, factor analysis, learning, and on principles of development. The customary "reports from laboratories" was held as a symposium on the first morning, with prepared summaries of certain studies in progress.

At the annual banquet held on Friday evening, Arthur G. Bills presided. Dean J. E. Le Rossignol, of the College of Business Administration, gave the address of welcome, livened with reminiscences of his study of psychology with G. Stanley Hall at Clark University and with Wundt and others at Leipzig. Walter B. Pillsbury delivered an anniversary address on the Nebraska Psychological Laboratory, relating particularly its early history under Dr. H. K. Wolfe. Edmund S. Conklin gave his presidential address on "The Status of Academic Psychology." In it he deplored some recent trends toward the dismemberment of departments of psychology in a few instances, even in universities where psychology has enjoyed a long and distinguished history. He issued a call for the type of leadership among psychologists that will gain respect, and for the minimizing of inconsequential differences which give outward impressions of inner lack of cohesion.

A special lecture on "Orientation of Organisms in Time and Space" was given by Karl Bühler as a contribution of the local chapter of Psi Chi, national honorary society in psychology, to the program on Saturday afternoon.

A resolution passed at the business meeting requested the secretary to thank the University of Nebraska, the host department, and the members of the local committee, composed of D. W. Dysinger, chairman, George R. Thornton, Warren

R. Baller, Leland H. Stott, Wilbur S. Gregory, Winona M. Perry, and Edmund E. Dudek, for providing facilities for an enjoyable meeting.

University of Nebraska J. P. GUILFORD

THE 1939 MEETING OF THE AMERICAN ASSOCIATION OF APPLIED PSYCHOLOGISTS

The American Association for Applied Psychology will hold its annual professional conference Friday through Sunday, December 1-3, 1939, in Washington, D.C. The Association had voted to meet with the International Congress of Psychotechnology should it be convened in America but that has proved impracticable. The By-Laws of the Association provide that "in so far as possible the Association shall coördinate its program with that of the American Psychological Association." It did not seem the part of wisdom, however, to take the annual meeting of so young an association so far from its center of population as would be necessary if it met this year with the American Psychological Association in San Francisco. The Applied Association voted, therefore, to meet for this year in the East at a time not in conflict with the American Psychological Association. There is no implication that the two associations shall continue to meet separately.

Ohio State University H. B. ENGLISH

Psychosomatic Medicine: Experimental and Clinical Studies

An introductory statement (pp. 3-5) in the first number (vol. 1, no. 1, January, 1939) of this new periodical explains that 'psychosomatic medicine' is not synonymous with psychiatry (a subject "concerned with the study and therapy of the disturbances of the mind whether these disturbances are the results of emotional experiences or of anatomical changes . . . of the central nervous system"), but "designates a method of approach to the problems of etiology and therapy." It is the "psychological approach into general medicine," which means, to the editors of the new magazine, "the understanding and the therapeutic utilization of the psychic component in the disease process and in the emotional relationship between physician and patient." That the psychoanalytic context shall not be confused with the old psychology, "so common in the last century," . . . "the medical reader" is warned that in "reading the expression 'psychosomatic medicine' he may be reminded of scientific or philosophical writings . . . as for example Wundt's 'Psychophysiologie' (sic) or 'Völkerpsychologie,' " books "devoted to scientific disciplines which were essentially nonexistent." [Who acquainted with the history even at second-hand would have blundered over Wundt's 'psychophysiologie,' would have thought the main psychologies of that time mere programs for the future and not "a report of facts already developed on the basis of observations," or would have said that the "main occupation of psychologists and philosophers consisted in writing introductions to scientific disciplines that did not yet exist."]

Over and above the exploitation of psychoanalytical concepts, methods, and theories, which the initial issue may well lead subscribers of the new journal to expect, the magic term 'totality' is obviously to be made fundamental. The division of "medical disciplines [why 'medical'?] into physiology, neurology, internal medicine, psychiatry and psychology" is said to be a mere convenience without biological validity. While the editors avow that there is "no logical distinction between 'mind and body,' mental and physical," they nevertheless demand that "psychological phe-

nomena should be studied in their psychological causality with intrinsically psychological methods and physiological phenomena in their physical causality with the methods of physics and chemistry." Powerful is the magic of the words 'unity' and 'psychosomatic totality'!

Fifteen years ago the present reviewer found the compound 'psycho-somatic' permitted in the larger English dictionaries as a formal adjective for mind-and-body but seldom if ever brought into psychological service. He used it systematically in *The Field of Psychology* (1924) to imply moments inherent in the psychological functions. It proved to be a bad concept, for it opened the way to psychical forces and agencies, entities which are never observed and had better not be assumed in scientific contexts. The term was presently withdrawn and condemned, with the remark that it failed "to avoid the intermixture of magical with natural causes, and it failed also to relieve psychology once and for all of the monstrous 'mind-body problem'" (*The New Field of Psychology*, 1934, p. vii). Soon, however, the word 'psychosomatic' began to reappear in medicine, in 'psychobiology,' and in those psychiatrical theories and cults which revive old speculations about the unconscious and about a vitalistic libido. Now that ancient pair 'body-and-mind,' made one by the abracadabra of 'unity,' is again current, its articulated pluralities artfully concealed under such phrases as 'the total organism,' 'the whole man,' 'the synoptic vision,' the 'he' or the 'she,' and 'psychosomatic medicine,' all phrases incident to a widespread and violent reaction against the analytic *many* and favorable to the primacy of the *one*. The violence of the movement, as well as its pendent character, is well indicated by the recent name of M. M. Black's book *The Pendulum Swings Back*.

Besides the introduction the first issue contains three writings by F. Alexander. The chief contents of the number fall under (1) a review of recent contributions on the hypothalamus and (2) a symposium on hypertension. W. R. Ingram writes a competent and useful review of the morphological and functional investigations into the hypothalamus, a review which is notably free from the speculative constructions which have multiplied of late in this difficult field. The final section, reserved for 'reviews, abstracts and correspondence,' will doubtless acquire substance and significance as the journal gets under way.

The concept of emotion as 'dynamically' understood by physiological and medical theorists seems destined to supply a major theme for the new periodical, although psychology (one of those 'invalid' division of the 'medical disciplines') takes its modest place along with psychoanalysis, internal medicine, neurology, physiology, psychiatry, comparative physiology, and pediatrics.

The new periodical is to maintain a mailing address at the building of the New York Academy of Medicine and to conduct its business through the sponsoring committee at the National Research Council in Washington.

Library of Congress, Washington MADISON BENTLEY

MONOGRAPHS OF THE SOCIETY FOR RESEARCH IN CHILD DEVELOPMENT

The Society for Research in Child Development has added this monographic series to its quarterly journal, *Child Development*, and its bi-monthly issues of *Abstracts and Bibliography* in the same field. Three volumes of 6 monographs each are now complete. They "consist primarily of detailed critical reviews and seriatim studies" in the wide field of child development, which is advertised to include

"anatomy and physical growth, physiology and biochemistry; nutrition; dentistry; pediatrics; physical health and disease; public health; mental hygiene and psychiatry; mental development and behavior; heredity, eugenics, evolution; education; sociology and economics." Of the first eighteen monographs, 14 are studies, reviews, and critiques in the large subject of physical and behavioral growth and development; 2 investigate mental aspects of childhood, 1 is devoted to personality of the child, and 2 are broad studies of methodology. Many of the studies are written by competent and experienced persons, who review and assay a large amount of existing research, chiefly American research. F. K. Shuttleworth and his associates have made further notable contributions to general somatic growth and sexual maturing; F. N. Freeman has recorded his evidence from repeated tests upon intellectual ability; N. Bayley has brought together a large amount of material on motor activities of the first three postnatal years, and J. W. Macfarlane has greatly aided all guides of the child by her wise and exhaustive survey (vol. 3, no. 6) of methods available for collecting and organizing fundamental data.

Those who have responsibly followed the checkered fortunes of the Society for Research in Child Development, of the antecedent committee of the National Research Council, and of the many agencies which operated under the Spelman Fund, will find high satisfaction in this monographic addition to the publications of the present Society. Let us hope that the Society's courageous efforts toward support and expansion will be sustained by all who profit from its competent labors.

Library of Congress, Washington, D.C. MADISON BENTLEY

Thomas Russell Garth: 1872-1939

Thomas Russell Garth, Professor of Experimental Psychology at the University of Denver, died unexpectedly as a result of an attack of asthma, on April 20, 1939. He was born in Paducah, Kentucky, on December 24, 1872. He received his Bachelor of Arts degree from the University of Denver in 1909, and his Master of Arts from the same institution in 1910. His work for the doctorate was carried on at Columbia University, and the degree of Doctor of Philosophy was conferred in 1917. After serving as Adjunct Professor of Psychology at the University of Texas, he was appointed to the faculty at the University of Denver in 1922, a position which he held until the time of his death.

Garth's earlier work included a study of the psychological effects of the hook worm, an investigation carried on by the Rockefeller Foundation in which he co-operated with E. K. Strong. After he came to Colorado his interest centered in the study of the American Indian. At first he drew attention to the differences between Indians and whites as measured by psychological techniques. In his studies of the full-blood and mixed-blood Indians, he found that, the greater the percentage of Indian blood, the lower the score made on intelligence tests. He also showed that, generally speaking, the Indian showed no appreciable superiority in sensory acuity over the white. Hampered by a lack of funds he made use of many small grants and drew heavily on his own means in order to carry on his studies. His students were always enthusiastic in their coöperation, and he said frequently that without them he could not have made a beginning in his work.

In 1935 he came in contact with several American Indian students who had been raised in white families and who showed a much higher degree of mental ability than he had observed before. Thus he determined to make a general study of foster Indian children raised in white families, and he spent his summers to find well over a hundred such cases. In a number of these instances he was able to test the siblings that had been reared in the Indian environment. His results showed that the mentality of foster Indian children in a white home is about equal to the mentality of the white members of the same home. Dr. Garth at the time of his death seemed to be approaching the hypothesis that the differences in mental ability between races are largely dependent on environment and the opportunity for racial advancement rather than on racial stock.

Dr. Garth was a member of the American Psychological Association and was President of the Rocky Mountain Branch in 1932 and 1933. Among the best known of his psychological books are *Race Psychology* and *Educational Psychology*. He recently published *The Life of Henry Augustus Buchtel*, former Chancellor of Denver University, an appreciative and well written biography.

GEORGE T. AVERY

Colorado State College

BOOK REVIEWS

Frustration and Aggression. By JOHN DOLLARD, LEONARD W. DOOB, NEAL E. MILLER, O. H. MOWRER, and ROBERT R. SEARS, with the collaboration of C. S. Ford, C. I. Hovland and R. T. Sollenberger. New Haven, Yale University Press (for the Institute of Human Relations), 1939. Pp. ix, 209.

This book promises more than it achieves. Introducing itself as a sample of coöperative research at Yale's Institute of Human Relations, it promises to illustrate the initial stages of an inductive-deductive method of science applied to this special problem of social psychology and to go about it with logical rigor. Ultimately it achieves five interesting essays which estimate and evaluate the variety, prevalence and inevitability of frustration in human society and, much less fully, of aggression. These essays attempt to present aggression as an immediate consequence of frustration and thus to support what is the regnant hypothesis of the book. Let us consider the essays first and the logic afterwards.

First there is Chapter 4, *Socialization in America,* which gives a genetic inventory of American frustrations from weaning to death. It is interesting to think of America as a Great Frustrator, and the list of frustrations is impressive. Here one learns something too about American aggressions, but the exposition neither strengthens nor weakens the common belief that frustration always leads to aggression. The discussion assumes this proposition; it does not demonstrate it.

The chapter on *Adolescence* and its frustrations (Chapter 5) is like a review of the literature in the *Psychological Bulletin.* It gives many facts in a rather uninteresting manner.

The best chapter in the book is *Criminality* (Chapter 6), a long essay which considers the dependence of criminality upon frustration with respect to twenty items, like educational status, intelligence, age, physical size, race, home conditions, drug addiction. The chief author here—the authors of the respective chapters are not specified although different authorship is indicated by different styles—is definitely trying to marshal the evidence that frustration leads to criminality. He is the more convincing because he finds doubtful eight of the twenty suggested conditions, and only twelve sure.

Almost as good is the chapter on frustration and aggression in modern America, Germany, Italy, and Russia (Chapter 7). Here it becomes clear how different are the frustrations of different cultures, how a nation may direct and displace the aggressions of its people, and thus how race prejudice, even though unreasonable, may stabilize a frustrated group. Although there is still no proving of the basic hypothesis, the description is plausible and convincing.

The last chapter takes the warlike and bold Ashanti and shows what frustrations are induced by their culture and how aggression apparently results. The Ashanti act as if they had planned their culture in the light of the major thesis of this book. They have, for instance, an annual cathartic festival in which for eight days anyone can tell anyone else, even the king, what he thinks of him!

So much for the essays. Now for the logic.

The book claims to be devoted to the consideration and the establishment or revision of the hypothesis "that the occurrence of aggressive behavior always pre-

supposes the existence of frustration and, contrariwise, that the existence of frustration always leads to some form of aggression" (p. 1). This statement is put forth as the "basic postulate" in a chapter of definitions. The authors are aware of the danger of circularity; they know that, if all consequents of frustration are labelled *aggression* simply for that reason, they will have begged the question. Thus they seek "independent" definitions. "Frustration is independently defined as *that condition which exists when a goal-response suffers interference.* Aggression is independently defined as *an act whose goal-response is injury to an organism* (or *organism-surrogate*)" (p. 11).

The meanings of the chief concepts are made clear in the following manner by their psychological relationships. (a) Goal-directed behavior must start from some cause, be initiated by some situation. This first term is called the *instigator,* a neutral term. (b) The *goal-response* to the instigation is activity that is terminated when the goal is attained. There can be no doubt that this is the valid definition of purpose. The goal is what terminates an activity that leads to it; the direction of activity toward the goal is nothing other than the fact that it will be terminated by the achievement of the goal. (c) Activities nevertheless come in time to lead fairly directly to goals because of the *reënforcing effect* (law of effect) of previous successful achievements. (d) "An interference with the occurrence of the instigated goal-response at its proper time in the behavior sequence is called a *frustration*" (p. 7). Good. "In order to say that a frustration exists, then, one must be able to specify two things: (1) that the organism could have been expected to perform certain acts, and (2) that these acts have been prevented from occurring" (p. 7). (e) In the face of interference with activity, there is most often a *substitute response* in which some similar goal (raisin pie when there is no mince) is reached and the goal-activity terminated or diminished. (f) If there is interference with the activity and no substitute response, then the frustration leads to *aggression:* the organism tries to injure another organism, or itself, or a surrogate. (g) Often the aggressive activity is *inhibited*—usually by punishment in organized society—and the aggression has to find some substitute. (h) Thus aggression may turn in upon the self, but *self-aggression* normally occurs only as a last resort since it is painful. Its occurrence usually means that all other avenues for aggression are strongly blocked. (i) The *displacement of aggression* and its discharge upon a surrogate may terminate the instigated activity, the whole process thus constituting *catharsis.*

Dr. May, the Director of the Institute, says in the *Foreword* that the authors are using a well known scientific procedure: seeing a *problem,* they set up an *hypothesis,* criticize it in the light of available data, refine it, and repeat the process. The ultimate goal is mathematical statement and the predictive value that it yields, but Dr. May does not pretend that the social sciences get so far nowadays.

Now let us agree to this definition of frustration and admit further that frustration is frequently followed by aggression and that aggression when blocked tends to be displaced upon some surrogate. So much fact is supported by common sense and a wealth of psychological evidence, and the five essays of this volume increase its plausibility. The book, however, is not content with such a loose generality. It says: if Frustration, then Aggression, always; and conversely. But what about the negative cases? Can not one find aggression without frustration? Frustration without aggression? Are there not alternative hypotheses to be examined? The history

of science is full of wrong hypotheses which were plausible enough and which were used for ages simply because they worked for a description of the phenomena, so long as the possible negative cases were not vigorously examined.

So we may ask: Is there not aggression without frustration? Now it can be argued that all mental life is aggressive, that striving (cf. McDougall) is exactly what it is to have a mind. In that case aggression could not always depend upon frustration unless frustration were universal too—as it might be if "divine discontent" is the essence of all motivation. This view of mental life is a very useful one. The organism is always striving for change; when it succeeds, it begins to strive for another change; when it fails, it persists and strives for a substitute; its persistence is noteworthy for it keeps on in spite of continued failure until it either achieves a vicarious success or is distracted. Aggression thus becomes the rule of life, but such aggression does not depend on frustration, although frustration depends on it.

Our authors, however, will object. They will remind us that aggression is defined so as to include only those activities that are injurious to another organism or its surrogate. To make injury the goal of aggression is definitely to limit the concept, but will such a limitation work? Is it aggression to strike a child? to spank a child? to spank a child for his own good? for your own good? for both his and your own good? Is it aggression to injure a child by giving him too much candy because you love him? Is it aggression to ride hard because you are angry, and then what is the surrogate object? My guess is that the authors are really appealing here to the outmoded introspective psychology of motivation, that for them activity is aggressive only when the subject *intends* harm, when he *hopes* for injury. That criterion is subjective and intuitive. It is not actual injury that is the goal of aggression because the aggression can be discharged without coming into direct contact with its object. Moreover, it is not clear what constitutes an adequate surrogate. If I denounce my enemy to the echoing hills, what have I injured? All in all it seems to me much clearer to say that aggression is any striving toward a goal, with the effort increased when there is interference until the interference is overcome or until the organism is redirected.

Now for our other question: Is there frustration without aggression? The answer would have to be *No* if all mental life is aggressive. On the other hand, if we try to keep to the narrow meaning of *aggression* which our authors had in mind, then there would seem to be exceptions to the Yale rule in the case of apathy and of distraction.

Lewin and associates (Patterns of aggressive behavior in experimentally created "social climates," *J. Soc. Psychol.*, 10, 1939, 271-299) have found that children, organized for collective endeavor under the restraints of an autocratic leadership, may experience frustration and react with aggression—often displaced aggression. On the other hand, they also find that, under certain conditions, the restrained children become apathetic and lose interest. In everyday life admission of defeat, withdrawal from the field, is common enough. What else is the meaning of *He shrugged it off?* Perhaps our authors will say that withdrawal from the field is the turning of the aggression in upon the organism's self, but surely that is nonsense. Do I never want food and put the thought aside without abusing myself? Of course, there are the traumatic frustrations which can not be put aside so easily, but there is nothing in this book to distinguish the simple interruptions of the goal-responses from dominating frus-

trations that will not be dismissed. The psychoanalysts think that fatigue is a symptom of repression. Is fatigue self-aggression or is it more like apathy?

Distraction is another normal escape from frustration. Our authors say that every frustration leads to aggression, although common sense would have it that the frustrated man is easily distracted. Suppose the young man, who had just been humiliated by the traffic officer, had then met his charming girl friend and she had started to admire him (p. 12), would his aggression (desire to injure the officer or his surrogate) necessarily have persisted? Have women never distracted men from aggression? Does a girl never recover from a frustrated love when her parents take her abroad? The truth would seem to be more nearly this: Activity tends to persist until the goal is achieved; interference heightens the activity; if progress toward the goal still fails, there is frustration, which results in (a) displacement of the activity to another goal, (b) apathy or (c) distraction of the subject to another activity. Do these authors really maintain that the activity of the salesman, who, having just achieved frustration in his attempt to make a sale, finds the building on fire, is in any discoverable degree dependent on the frustration, that there is no such thing as distraction and only substitute response?

Such a view is too universal—that all distraction is displaced aggression and all apathy is self-aggression. Then it would be only after perfect attainment of a goal or after complete catharsis that unaggressive activity could take place. It would become conceivable that the neonate, frustrated by his birth, would spend the rest of his life working out that perpetually reënforced original aggression!

Dr. May says that these five authors found the coöperation of writing this book together frustrating. In fact he makes his statement general: "coöperative work of an intellectual nature which penetrates beyond the level of superficial collaboration is itself frustrating" (p. vii). That statement I challenge. Coöperative work can occur in serious undertakings among intellectuals without frustration (a) when there is accepted leadership, (b) when there is clear division of function and the acceptance of a separate leader for each function, and (c) when there is so strong a desire for coöperation that the individual experiences a sense of achievement in surrendering his initial thesis for the ultimate synthesis. If the Institute wants coöperation it might have it on these terms, but my own advice would be for it to stick to competent individualists and to let coöperation go except when it happens spontaneously without frustration.

In a word, then, this book enlightens us with many interesting facts about frustration and aggression in social life, it seduces us (almost) with a plausible but defective logic of the universal concomitance of frustration and aggression, and it warns us to scrutinize the omens with great care before we venture upon coöperation.

Harvard University EDWIN G. BORING

Foundations of Education Psychology: Nature's Gifts to Man. By PETER SANDIFORD. New York, Longmans, Green & Co., 1938. Pp. xv, 464.

The reviewer had mixed feelings as he read this work which the author tells us in his preface is the first of a series designed to give a general survey of the foundations of education psychology for advanced students of education. The field covered by this first volume, he tells us, is approximately that which Thorndike called, "The

Original Nature of Man." A second volume, he promises, will deal with the "Psychology of Learning."

The writing of an up-to-date two or three volume treatise on educational psychology which will supplant Thorndike's out-of-date three volumes is a much needed work. It would seem, however, that Professor Sandiford has not gotten off to an exceptionally good start in this task, nor has he succeeded in writing a thoroughly acceptable book for the purpose stated in the preface. There are a number of prominent lacunae occurring in the discussion of certain topics, which omissions are really inexcusable in a book written for graduate students. Before pointing out some of these omissions let us point out some features of the treatment which pleased the reviewer.

In his first chapter on, "The Nature of Educational Psychology," there is an especially valuable summary, for the space allotted to it, of the main schools of psychology. The criticism of Gestalt and Organismic psychology and of certain fadish ideas and procedures which *some* progressive educators have attempted to justify by reference to these schools of psychology is particularly apt. The chapter on "Personality" is a valuable summary. Throughout the book many valuable researches are summarized in a helpful way with the aid of tables and charts; and the significance of the results well integrated. The reviewer would have liked more discussion of the educational applications of the various findings; but the author has chosen to leave that for the reader to supply in any large quantity.

In his discussion of heredity and environment it is unfortunate that Newman, Freeman, and Holzinger's study of twins is accepted at face value and in toto. The publishers have assured the reviewer that Sandiford's book went to press before McNemar's significant criticism of the Newman, Freeman, Holzinger study appeared in the April, 1938, *Psychological Bulletin*. It is unfortunate that such was the case. Otherwise a publicizing and approving of a study whose concluding chapter does not seem to follow from a careful analysis of what lies between the first and last chapters might have been avoided.

In the discussion of the conditioning of the emotions he stresses Watson's first experiment on conditioning fear in an infant and fails to mention. Bregman's work at Columbia and the observation of individual infants by two or three others, which show the difficulty of conditioning emotional responses, under certain conditions, by the technique which Watson used. The discussion of the training of the emotions overemphasizes the re-conditioning technique for removing fear responses and fails to mention the recent findings of Jersild and Holmes which suggest other techniques as more practical. The latter study is mentioned by title only, some pages later; but the findings and their significance are ignored.

The section which this reviewer found most unsatisfactory was the one on "Intelligence: Its Nature and Measurement." The first two lines read, "Intelligence is innate. Studies of heredity show that intelligence is innate, that it actually has its basis in the genes." What is this innate intelligence? A few pages later he tells us that the commonest features he finds in scores of definitions in his collection are, "(1) the ability to adapt oneself to novel situations; (2) the ability to carry on the higher processes of thought, especially abstract thinking; and (3) the ability to learn" (p. 337). Further on, regarding Spearman's theory of intelligence he says, "For this two-factor theory, the evidence supplied by Spearman and his students over a period of two decades seems to be overwhelming. . . . It should also be

noted that the new techniques of factor analysis, instituted by Thurstone, Hotelling and others, lead to the discovery of factors such as language ability, motor ability, etc., when intelligence and other forms of tests are investigated. . . . Yet after examining the evidence, the balance is undoubtedly on the side of a unitary conception of intelligence" (p. 343-4). In an earlier section of the book twenty pages are devoted to the theory of the gene and the mechanisms of heredity. The reviewer, however, fails to find any evidence there, from psychologist or geneticist, nor any evidence in the author's discussion of family history and correlation studies, that would support the statement with which this section on intelligence begins; and which implies that a "mystical" g and the three "commonest features" of scores of definitions are the innate trait of intelligence. (He cannot help reminding the author of the discussion of the scientific method and principles, especially that of parsimony, which is so well presented in the introduction.) Among psychologists there is all too much confusion of this "mystical" intelligence with the "intelligence" which is what intelligence tests measure. What we can really talk about in any than a speculative way, are the test scores themselves and the effects of hereditary and environmental factors upon these scores. Discussions for the educator should be confined largely to this approach. Too much damage has already been done by the educator's misinterpretation of the psychologist's IQ as something accurately determined from one paper test, 'mystical,' innate, and unchangeable.

Even though he accepts Spearman's theory on the nature of intelligence that does not excuse him, in a book "for advanced students of education," for omitting, in the discussion of intelligence, any specific mention or even bibliographical reference of Thurstone's criticism of Spearman, appearing in 1934, nor for omitting Thurstone's more recent theory of seven primary abilities which was reported in print in 1936 and more extensively this past year.

The section on intelligence should have contained a much more critical discussion of the constancy of the IQ, mentioning the variations between individual and group tests and between various group tests. A much more extensive and critical discussion of the environmental influences upon IQ should have been given. In this connection, even though the author may not accept the conclusions, he is bound in a book of this type to mention the studies of this problem that have been coming from Stoddard and Wellman and their associates since 1932 and which culminated in the study of "Mental Growth from Preschool to College," appearing in December 1937.

In discussing intelligence tests for preschool children he fails to mention at all the Minnesota Preschool Scale; and although mentioning the Revised Stanford-Binet in other connections, fails to mention it as a test for preschool children. This is most unfortunate since they both appear to be much better tests of what we usually test by intelligence tests than the Merrill-Palmer Scale which he does go to the trouble to describe.

There is running through the book a careless tendency to omit, at times, the dates from the volume and page references to articles. This provoking omission is certainly not to be condoned.

It would seem that the text assumes a good bit of biological and physiological background which many education students may not have or be able to get readily. I wonder if we really should force this type of material upon them, or whether we should instead, give them the implications of these studies, with the references to the studies themselves. Then the student with the background and interest may look

into them for himself. Rather than twenty pages on the mechanism of hereditary the reviewer would have liked to have found twenty pages discussing the best opinions of biologists and environmentalists concerning the contribution which eugenics and euthenics can be expected to make to the welfare of society.

Sandiford has written a valuable book. The advanced student whose knowledge enables him to recognize the lacunae will profit much from reading it. It would be too bad, however, to put it into the hands of the student just beginning his graduate work in educational psychology unless it is studied with warnings and critical guidance from the instructor.

Pennsylvania State College EDWARD B. VAN ORMER

The Achievement of Personality. By GRACE STUART. New York, Macmillan Co., 1938. Pp. 192.

To the typical American reader, including college students, the title of Grace Stuart's book will be misleading. To the psychoanalytically and theologically uninitiated it will no doubt be unintelligible. The title, *The Achievement of Personality,* probably suggests to most people a work devoted to a definition of the concept of personality and to intelligent ways and means of developing those desirable qualities that make for social acceptance and personal security.

Actually the book is concerned with a commendably ardent discussion of what is apparently believed to be the only logical avenue to the rejuvenation of man and the revamping of society. The author proposes an "air tight" philosophy, the essence of which is that to be happy and to resist the ravages of neuroticism man must be imbued with the spirit of love and must in turn be beloved by both his fellows and his God. This is of course a basic truism that, in its more prosaic and psychologically objective aspects, has been known to clinical workers for some time. The principle has not, however, been any too widely understood and applied by a frequently faltering society which could certainly benefit by its exemplification.

The impression one gets from reading this book is that it is more a psychology of religion than a psychology of personality, at least in the functional sense. It is purely subjective, quoting only other subjective excerpts from the writings of such countrymen (English) of the author's as McDougall, Hadfield, Brown, and Jacks, from American therapists such as Hocking and Pratt, and from the three psychoanalysts, Freud, Adler, and Jung. The book is really an integration of subjective psychology, Christian philosophy, deistic religion, and psychoanalysis, with strong accents on the latter two. Objective data of the more earthly variety as developed by Gesell, by Hartshorne and May, and by J. B. Watson are conspicuous by their absence. This omission was obviously in the intent of the author but it means that her book is decidedly abstract, idealistic, and far removed from the concrete channels of everyday living.

A brief résumé of some of the author's "pegs" on which she hangs her concept of the union of egoism and altruism as the answer to man's quest for happiness should be enlightening to the reader. First of all man is actuated by a dynamic "life force" which partakes of the nature of Schopenhauer's "will to live," Bergson's "Elan Vital," McDougall's "instincts," and the psychoanalysists' "libido." This life force is personalistic rather than mechanistic since it constantly reaches out to get

the things that the human organism needs. The personality becomes organized by developing emotional dispositions in connection with objects and people in the environment as per Shand's law of "sentiment-formation." Personality is a "seed" or "soul" which is to be so conditioned and nourished that the sentiments accruing will favor the Good, the Beautiful, and the True. The psychophysical life force (and here enters dualism) is to be given the "out-turned" direction that makes for magnanimous love and that obviates neurotic self-concern.

But the personality has driving needs which must be reasonably well met if it is to resist disintegration. To the author these needs are summed up, with some reservations and amplifications, in Freud's "need for love," Adler's "need for significance" (will to power), and Jung's "need for security." The need for love and the need for significance are said to be somewhat synonymous but the need for security as propounded by Jung is broader and more fundamental in that it introduces the religious principle that man thrives best when he struggles valiantly through adversity and tribulation toward his precious goal of self-realization through sacrifice of self-concern. Neuroses can only result when love, recognition, and security are not realized. But love, recognition, and security cannot be sought directly, they must be won by way of an outflow of love for humanity.

Man's energy, which is originally inclined to satisfy the selfish desire for pleasure (Freud), power (Adler) and indolence (Jung), must eventually come to flow *out* toward sacrifice and social altruism. This means moral conflict for the growing personality. The noble ends of living are best but the immature self-desires are hard to give up. If and when the sacrifice is completely made sin, cruelty, war, prostitution, and other such miseries will fade away like a bad dream. Unequivocal love with its righteous behavior is entirely possible when man sets his affections on "lovable objects."

The author speaks of God rather gingerly. There may not be a God but psychologically man needs him. Psychotherapists should not expose their doubt on this point lest the therapeutic value of "God belief" be lost for their clients. But be God as he may, man's most imperative and irreducible need is an environment which provides him with love and in which he can love others. Finally, a world in which man must struggle for material and spiritual things is good for him, but a situation in which he cannot be, or is not, loved will inevitably lead to his destruction. Personality is strictly relational. To live man must be a lover and he must be beloved. This is the balance that good adjustment demands.

In terms of her purposes and her philosophy the author has certainly expressed herself logically and adequately. Each of her chapters leads naturally to the climax presented at the very close of the book. However, when one considers the dualisms, the "mind" and "soul" implications, the emphasis on instincts and psychoanalytic concepts, and the neglect of modern descriptions of concrete personality organization and development, it seems to the reviewer that the author is going through a stage of development in the personality field that characterized most American psychologists some fifteen or twenty years ago. Perhaps Mrs. Stuart is aware of this fact but wishes to eschew the objective monism which seems to be so dear to our temper.

University of Southern California LOUIS P. THORPE

Methoden der Psychologie des höheren Gefühlslebens. By G. STÖRRING. Berlin and Wien, Urban & Schwarzenberg, 1938. Pp. viii, 522.

The present volume was originally a contribution to Abderhalden's encyclopaedic handbook on methods of biological investigation, but is now issued as a separate work in order that it may be available to wider circles. It is the first work in any language which attempts to assemble, explain, and evaluate, all of the methods which have been employed in the investigation of feelings, emotions, and effective processes in general, both as analytic phenomena and in their relation to other psychological processes.

For the execution of this task the author is particularly well qualified, because of his distinguished contributions in this field, as well as his lifelong study of the relations between psychology and psychopathology. Indeed as early as 1900, Professor Störring produced the first systematic treatise on the significance of abnormal mental phenomena for normal psychology, a work translated into English in 1907, by Thomas Loveday, under the title of "Mental Pathology in Its Relation to Normal Psychology."

The present work on the study of affective processes of a higher order, is at once encyclopaedic and systematic. This is the happy result of a comprehensive aim, an admirable mastery of materials, and a gift for logical organization. It is therefore possible to study any particular section with profit, apart from the whole, although the book is organized in strict accord with a definite system of affective psychology.

The author divides the total field of affective psychology into general and special disciplines, devoting the larger portion of his work to the general field. In pursuing investigations of this kind, four methods have been employed; *viz.* (1) the psychopathological method, (2) the experimental method, (3) the introspective method and (4) the method of folk psychology. Of these the psychopathological method is fundamental according to the author. The other methods, however, are of great importance as independent means of confirming or illuminating the results of the basic method, while in some situations they offer the only means of discovering relevant affective elements and their laws. This is especially the case with the method of folk psychology which is indispensable in the investigation of the complex affective processes involved in ethical and esthetic situations, and indeed in all situations where human values are in question. It is notable that Dr. Störring true to the tradition of the Great Age of German psychology of which he is one of the last representatives, does not disdain the introspective method. He devotes more than one hundred pages to its uses and results, and suggests that affective consciousness is a form and energy which may find expression in overt behavior of profound significance to the individual and society. The direct study of such affective consciousness, by introspective methods, must be indispensable to any really constructive human psychology.

The first and longer section of the work is concerned with the general theory of affective processes and comprises three sections dealing successively with the nature and manipulation of the psychopathological, experimental, and introspective methods of investigating affective processes of all degrees of complexity and consciousness.

Full justice is done to the recent contributions of psychoanalysis and allied schools, while none of the classical psychologists or traditional schools are neglected.

The second part of the book treats of the special theory of affective processes, and presents a discussion of feelings of value and specific emotional judgments of worth as these occur in ethics and esthetics. The second section of the discussion presents a schema of the investigation of ethical value judgments by means of the techniques of folk psychology, and the work closes with a chapter on the logic of valuation in the normative sciences.

In his treatment of an extensive literature, Dr. Störring exhibits a fine historical perspective and a rare sense of the relative importance of different contributions to affective psychology. He summarizes briefly the methodological contributions of all important writers from every school of psychology whether old or new. Without wasting time in polemics, he indicates the value and limitations of the work and whatever contradictions may exist with the works of others. He is apparently considerably impressed by certain correlations between Ach's recent findings in the investigating emotional objectivation, and the typological theories of C. G. Jung. Capacity for emotional objectivation (Objectionsfähigkeit) appears to be correlated with extraversion and increased suggestibility, while a weaker capacity for emotional objectivation seems connected with introversion, reduced suggestibility, and psychic resistance. If true, this generalization may justify the theory of an emotional basis for civilization as proposed by certain historians and social psychologists.

Although the author gives precise references to all authorities in footnotes, a general bibliography would add to the value of this splendid work, which merits translation into other languages, as a general guide to investigators of the affective life of man.

University of Cincinnati CHARLES M. DISERENS

"Truth" as Conceived by Those Who Are not Professional Philosophers. By ARNE NESS. Oslo, I Kommissjon Hos Jacob Dybwad, 1938. Pp. 178.

The writer of this dissertation in English from the Norwegian Scientific Academy in Oslo is motivated by the laudable desire to find out exactly what those who are not professional philosophers mean by "truth." He is impressed by the glib way in which philosophers speculate as to the plain man's view of truth without any empirical evidence to back up their formulations thereof. "Why," he asks, "is observation systematically neglected" in this matter, and "sweeping, vague, general statements" offered instead? Feeling convinced that it would be of value to know definitely what non-philosophers really do understand by truth, our author proposes to determine this in a scientific way, by constructing situations calculated to arouse verbal reactions adapted to the solution of this problem.

The method used is that of the questionary. Some 250 "test-persons" were selected for examination, and about 6000 hours were expended in the prosecution of the study. The subjects varied in age from 7 to over 30 years, and in educational background correspondingly, and there was a slight preponderance of males among them. But the author is not at all clear or systematic in his presentation of the essential data. He employs also a quite outlandish system of abbreviations, and the list of them which appears at the beginning of the book is incomplete and not particularly helpful.

The central question under investigation is the following: "What is the common characteristic of what is true?" But a number of other questions are also propounded,

such as these: "Give an example of something that is true." "Is there something absolutely true?" "What is a fact?" "Mention something you know with certainty." "How do you explain that you know these things with certainty?"

The answers are about as diverse as there are subjects to affirm them, but they are grouped by the author in a number of ways too numerous for us to recount here. Typical answers are such as these: Truth is "agreement with reality," "something that happened," "that which can be proved," "that which cannot be doubted," "what is arrived at by using one's senses," "something that must be thus and thus." Many of the subjects are skeptical as to whether truth has any common characteristics at all.

The author has not formulated his conclusions in any systematic fashion, and all we seem justified in saying about them is that whatever results *do* issue from these elaborate questionnaries are of a purely tentative, not to say negative, character. Mr. Ness himself estimates that 30,000 hours would be necessary for accomplishing anything of positive value—five times the paltry 6000 hours already actually expended by him. "No marked, real correlation" is found between the character of the answers and the age or educational status of the subjects—or between any other two factors, so far as we can discover.

As to the fundamental question, What is the notion of truth common to all non-philosophers? the only possible answer, if our investigator is right, is this: There is no such common notion of truth among non-philosophers, any more than there is agreement as to the nature of truth among philosophers themselves. On these and like matters, indeed, our author finds little difference between these two groups of persons: philosophers, we are told, like to stress their own cleverness by calling non-philosophers "stupid," but after all there is not much to choose between them. Nor, finally, is there any more justification for calling the common sense view of truth "realistic" than there is for calling it non-realistic.

This book, in the opinion of the reviewer, suffers from an excess of complex methodology quite out of proportion to the meager results finally obtained. The author tries desperately hard to be "scientific," but seems to have little to show for it when all is done and recorded. His English is not always above reproach, but it is, nevertheless, generally speaking commendably free from solecisms, and there are fewer misprints than are usually found in works of this character.

Western Reserve University JARED S. MOORE

Analysis of Parergasia. By GLADYS C. TERRY and THOMAS A. C. RENNIE. New York, Nervous and Mental Disease Monograph, 1938. Pp. 202.

As those acquainted with the work of Adolph Meyer and the psychobiological school will recall, *parergasia* has been proposed as a more plastic term for schizophrenia. This monograph proceeds from the concept of schizophrenia as dependent on a special constitution and personality characterized by habit disorganization. The authors emphasize a genetic-dynamic understanding of the patient through a study of his total life record and of the factors entering at various integrative levels. By such an approach alone may we comprehend the chaotic schizophrenic reactions and their meaning to the individual patient himself. In contrast to a fatalistic acceptance, the focus is on potential re-synthesis "in recreating the genesis and dynamics of the life record" and by management of developments in the psychosis itself.

The material analyzed comprises 77 case histories selected from a preliminary survey of approximately 250 cases diagnosed as parergasia at the Henry Phipps Psychiatric Clinic. The selection was made for the purpose of singling out items that have proved significant in the life records of the parergasic group. A key to these individually coded and charted histories presents an extensive list of topics under nine major headings. Seven of these headings deal with items of import in the pre-psychotic lives, the other two with features of the psychotic development and outcome. In any study of parergasia this key should be of crucial value; particularly when one realizes the time element involved in assembling the material, the completeness in telling items, and the fact that the services of a large, varied staff were utilized.

For the purpose of differential diagnosis the monograph offers six contrasting clinical reaction-patterns: aggressive, conversion, passive, deliroid, affective, and defeatist with their respective sub-groupings and descriptions. Each pattern is illustrated with case material. It is a credit to the scientific prudence of the authors that nowhere do they attempt rigid categories or unwarranted statistical refinements. Frequencies of items sifting out from the material studied are adequately tabulated under topics taken from the key. These tables serve as a frame of reference for subsequent discussion.

According to a basic formula psychobiological happenings are conceived as *somatogenic, endogenic, exogenic,* and *neurogenic.* The resultant analytic summary of this heterogeneous material suggests the following as characteristic of the parergasic group: (1) *in the organism,* we find striking incidence of instabilities and oddities in parents and siblings; asthenic habitus, dysplaic features; early unstable or inadequate behavior, fears, solitary trends, poorly directed aggression; (2) *in the subject,* disturbances in menstrual, gastrointestinal, cardiovascular, and skin-functions; dependence upon parents, poorly organized sex trends, indecision in love and martial problems; exaggerated need for approval, feelings of inadequacy, shyness, interests limited, self-centered and lacking in spontaneity; (3) *in the situation,* unstable parents, particularly mothers, over-attachment to the mother; domineering or over-protective family, unfavorable contrasts with more successful siblings; rarely achieved extra-familial adjustments, sensivity to opinion, incongruous appearance of indifference; increasing asocial trends; (4) *in the brain,* events inducing convictions of inadequacy, conflictual strivings, loss of inhibition, overreaching goals, fantasies, suspicions, uncritical thinking with acceptance of delusions and hallucinations, loss of planning, scattering.

In short, the parergasic personality is fundamentally a dependent and passive personality whose basic developmental features show an insufficiency reaction making for discrepancy. The life record manifests much *Sturm und Drang,* with cataclysms along the way determining the direction of the upheaval. In adolescence particularly attempts to fuse sexual and social needs are poorly integrated. For mental hygiene the implications are clear-cut and imperative. Unfavorable familial and extra-familial factors, early manifestations of sensitiveness and unfavorable comparisons must be recognized and guarded against. Thinking in the concrete is to be encouraged, as are modification of goals where necessary, desensitization of unfortunate sex experiences, and expansion of ego-centric interests. Alove all, security and understanding must be provided.

Although written primarily for a psychiatric audience *Analysis of Parergasia* will generously repay the serious student of abnormal psychology and mental hygiene. The point of view is constructive, never fatalistic. Throughout the material is handled in an orderly fashion; the cards are allowed to fall which way they will. The psychobiological approach developed by Adolph Meyer is reëmphasized and sustained. Wherever interpretations are made a high degree of empiricism and a will to see the interdependence of dynamic factors in the life history go hand in hand.

University of Maryland WALTER M. SPARKS

Modes of Thought. By ALFRED NORTH WHITEHEAD. New York, Macmillan Co., 1938. Pp. viii, 241.

Professor Whitehead's latest volume is a collection of chapters from various sources: the first six, a series of lectures delivered at Wellesley College in 1937-1938, on the general subjects of "Creative Impulse" and "Activity," and now constituting Parts I and II of our book; then two lectures on "Nature and Life," delivered four years earlier at the University of Chicago, and already published in separate form by the Press of that University; and a final epilogue on "The Aim of Philosophy," adapted from an informal address delivered to Harvard and Radcliffe students in 1935. In consequence of this method of composition, it is hard to trace any single thread of continuous thought running through the entire book, and the already-published lectures on "Nature and Life" are certainly clearer and of more general interest than the portions that are new. If one is pressed, however, to say what is the main purport of the book one might assert that *Modes of Thought* is a study of dualities, of pairs of contrasted ideas which must be joined together in any adequate world-view—such ideas as matter-of-fact and importance, particularity and generality, variety of detail and unity of significance, temporal process and eternal form, the actual and the potential. The world, Professor Whitehead seems here to tell us, is a vast conglomerate of constantly changing processes, ever pushing ahead to newer and newer goals; and it is the function of intelligence, not so much to guide the direction of these processes, as to choose between the more and the less significant ones, and to see them all in their proper perspective and in their relation to one another.

As of special interest to the psychologist, the following points stressed by the author may be noted: (1) the distinctive nature of man; (2) the importance of consciousness; (3) the function of language; (4) the nature of understanding; and (5) the superficiality of sensory experience. I will briefly elaborate each of these points. (1) Human mentality differs from that of the lower animals chiefly in its twofold capacity to conceive ideal possibilities and to understand the structure and formal relations of things: "to be human requires the study of structure, to be animal merely requires its enjoyment" (p. 105). (2) Consciousness is not an otiose accompaniment of bodily functioning, and still less a mere synonym for cerebrally initiated behavior, but a true "agent of stimulation" (p. 44), an emphasizing and selecting activity in the affairs of life. (3) Language has the function not only of communicating one's experiences to others, but also of expressing one's own past to one's present: it "is not the essence of thought," though without it thought would be "gravely limited" (p. 49). (4) "Understanding is not primarily based on inference," but rather is a matter of "self-evidence," though inference is

a "means for the attainment of such understanding as we can achieve" (p. 69): the function of understanding is the twofold one of seeing things in their unity and in their relations to other things. Finally (5) we call attention to Professor Whitehead's emphasis on the superficial character of sense-experience as revelatory of the nature of reality, notwithstanding its prominence in our conscious lives and its practical importance as a stimulus of activity.

The behaviorist, we see, will gain little consolation from the reading of this volume. If my exposition of its contents seems somewhat scrappy, this is a reflection, unless I am very wrong in my interpretation, of the disconnected structure of the work. As previously intimated, only Part III, containing the two Chicago lectures on *Nature and Life,* reveals any very clear unity and continuity of theme, and the book as a whole is really two (or even, perhaps, three) books and an epilogue. The final chapter constitutes a quite charming *envoi,* but it is hardly more.

Western Reserve University JARED S. MOORE

Psychological Foundations of Personality. By LOUIS P. THORPE. New York, McGraw-Hill Book Co., 1938. Pp. xvi, 593.

It is a rare week that does not bring a publisher's circular heralding the latest and most scientific (and therefore the best) book in this or that special field of psychology. Not a few of the brethren seem willing to affix their rubber stamps of approval to these circulars, though most of the books so certified, as the reader is disheartened to discover, are dilutions of dilutions of old wine. This *camaraderie* among psychologists does a service, possibly, to the cause of humanitarianism, but a disservice, certainly, to the cause of psychology. Perhaps the universities, pressing the notion that 'productive scholarship' should be the one criterion for preferment, are responsible. Anyhow the truth is there are too many undifferentiated duplicates on the market; and in my opinion an academic man should be rewarded for *not* adding to the surplus. (Just as the Government has seen fit to pay farmers for not raising hogs.) Having recently committed the offense of writing a book myself, I am entitled to express this view—though not without a blush.

As it happens there are not so many academic books whose title includes the word "personality," and one hopes for the best in opening a volume which the author describes as an attempt "to organize, to refine, and occasionally to re-interpret the many references to personality matters" ("matters" here makes one bristle) that have appeared in articles and books on other topics. Unhappily it is exactly in these respects—upon which the success of a book that offers no new material must hang —that the attempt collapses. Beyond the headings of the different sections, this reader, at least, failed to find any appreciable organization, refinement or re-interpretation. The author has been as democratic generally as a Salvation Army-man, opening his portals to every theory or opinion, employed or unemployed, that happened to be straggling along the highway.

Vowing his allegiance to objective scientific psychology, Dr. Thorpe sets out to treat personality genetically, beginning with the superior amateurs, Mendel and Galton, and ending with a horde of inferior professionals. In the name of science the instinct theory is ejected at the front door, but in a moment its brother, the drive (or need) theory, is admitted at the rear. There is a full, though loose, discussion of the concept of traits, to which the author gives wavering assent; and a

fairly complete account of paper and pencil tests of character and temperament. As usual the intelligence quotient receives prolonged consideration. Finally we come to the "wholesome personality."

The author of *Psychological Foundations of Personality*, being a professor of education as well as a practical therapist, is quite naturally interested in the "educability of personality;" and his book is advertised as a "handbook on the craft of inspired and purposive teaching." But there is no evidence that Professor Thorpe believes that "inspired" education and therapy must necessarily be based upon a philosophy, or system of values: matters which are still outside the province of objective scientific psychology. One must be honest about one's inspiration, and not use the masquerade of scientist while implicitly affirming moral preferences, personal or conventional. For one thing it leads to superficiality, which can be as misleading and lethal to the human spirit, as the worst fumblings of depth psychology. Perhaps Man is not such a riddle as we supposed. Perhaps the grievous mental ills of which he complains today are manageable. But only an irreclaimable optimist will believe that he can be fathomed by the unrelated categories or cured by the shotgun prescriptions which constitute this book.

Harvard University H. A. MURRAY

The Speech Personality. By ELWOOD MURRAY. New York, J. B. Lippincott Co., 1937. Pp. xii, 517.

This is a textbook and guidance manual for class and individual instruction in speech at the college level. A chapter is devoted to each of the following aspects of speech: (1) mental objectivity; (2) poise and emotional control; (3) spontaneity and emotional responsiveness; (4) audibility and distinctness; (5) vocal quality; (6) physical bearing; (7) gestures; (8) vocal flexibility; (9) enunciation, articulation, and pronunciation; (10) oral grammar and diction. These characteristics are classified as the *grosser speech skills*. They are antecedent to the development of the more refined speech techniques. Briefly, a speaker may be said to be mature and possessed of refined speech techniques when he is capable of leadership in speech situations and, also, when he is, by virtue of his speech skills, capable of making creative contributions to the social order.

The book appears to be a substantial contribution to the field of speech correction. The facts are presented clearly, the instructions and assignments for speech improvement are specific, and the practice exercises are well selected. A significant amount of practical, usable information is to be found in the chapter on enunciation, articulation, and pronunciation. The appendices contain outline sheets for history taking, personality improvement, hearing test, and voice analysis. In general, the more relevant aspects of speech correction are either included or conveniently cited so that the reader may have ready access to the sources. Although the book was mainly intended for class instruction and individual guidance, the clinical psychologist may find it of value in adult psychotherapeutic work whhere a speech defect is symptomatic of personality distortion. Unfortunately, its serviceability in such a direction is seriously handicapped by inadequate treatment of examination methods, an almost total neglect of clinical procedures, devoting but scant space to the specific speech disabilities.

It is our opinion that the title is not appropriate to the topic. It is neither easily

nor correctly defined, nor is it likely to be accepted as convenient usage among those in the field of speech therapy, especially if they be psychologically oriented. According to Dr. Murray, the speech personality "is the personality both as built up and as expressed through speech." If there is a "speech personality" then why not an "athletic personality," a "singing personality," a "business personality," etc. The term is loosely used according to accepted psychological terminology. Its generic significance is lost in its specific applications.

Dr. Murray's opinion that "speech development parallels personality development" is necessarily hypothetical. Certainly there are too many exceptions to this generalization. There is, moreover, no clear-cut experimental evidence to support the statement. Under the circumstances, it might have been advisable to minimize the point of view on personality development and to have referred the reader to the accepted standard works on personality and adjustment.

The factors which make this book a good text are possibly responsible for preventing it from becoming a relatively permanent contribution to the field of speech correction. The style, which is not always felicitous, is dogmatic. There appears to be ever present the implication that the student is being instructed and that the instructor's opinion is *fiat*. This proves to be, as a matter of fact, a really good method of presentation: the reader is rather readily convinced, especially since the words carry an inspirational enthusiasm.

Rochester, New York ANTHONY J. MITRANO

Experimental Foundations of General Psychology. By WILLARD S. VALENTINE. New York, Farrar & Rinehart, 1938. Pp. xviii, 377.

This book is another offering in aid of the growing tendency to introduce beginning students to psychology by way of its experimental literature. The value of such a digest rests upon its general agreement with a particular user's own aims and problems in teaching his elementary course. The following quotations from the three introductory statements by the author indicate his plan of presentation. "In the selection of the material that appears here my criterion has been its utility to a young person who in all likelihood will have only one course in Psychology." Appeal to the student has been kept in mind to the end that "some relatively trivial experiments were included because they have proved over a period of years to be of interest to beginners." Definitions are avoided but the author is obviously concerned with factors that determine a person's conduct and these are roughly divided into "(1) a person's heredity," "(2) a person's biography" and "(3) a person's surroundings." This classification largely determines the interpretation of the various selections.

The body of the text consists almost wholly of re-written descriptions of experiments and interpretations of the results of investigations, with occasional excerpts from original articles. The selections embrace a broad conception of psychology as the study of behavior and are collected under such customary chapter headings of a textbook of psychology as Emotion, Perception, Learning, Remembering, Reasoning and Intelligence. Considerable space is given also to borderline topics such as Inheritance, Drives, Market Research, Hypnosis, Conflict. There is no attempt to be systematic or complete in the treatment of the whole field or of particular topics. The author states at the beginning that the student is not expected to get a "rigorous

system of psychology," nor to be prepared "for an advanced course in any special branch of psychology." He will be interested at times and informed on many points psychological and otherwise. If such is the aim of an elementary course, the book answers that purpose. The material is well presented and sufficiently illustrated so that many instructors will find it useful supplementary reading.

Hobart College FORREST L. DIMMICK

The Rediscovery of Man. By HENRY C. LINK. New York, Macmillan Co., 1938. Pp. ix, 257.

This book portrays man as potential master of himself and his environment. In it the reader will find an extension of many of the ideas advanced in the author's *Return to Religion*. Personality is defined as "the extent to which the individual has developed habits and skills which interest and serve others." Link's measurement of it is predicated on four basic components: Habits of Self-Determination, Social Initiative, Economic Self-Determination, and Adjustment to the Opposite Sex. Within this frame of reference the author vigorously attacks a mechanistic concept of modern man as the helpless victim of forces revealed by the physical and social sciences. In his treatment he neglects none of the pluri-dimensional settings of personality be they individual, economic, social, recreational, religious, or political.

Many psychologists will probably neither approve nor agree with some of the statements contained in this book. They may object to such over-simplifications as: "Psychologists have found, however, that personality, far from being intangible, consists of definite habits and skills. These habits and skills can be acquired in the same way that people acquire the habits of reading and writing, that is, by practice and training." They may further protest at finding the solution to universal problems in development by the individual of correct habits. They will find unwarranted generalizations, will not be slow to point out that integrated personalities may greatly suffer from mis-directed attempts toward socialization, and will strongly resent the frequent labelling of personal opinion as fact discovered by Psychology.

The author, however, addresses laymen and not psychologists. Chapter I appeared last spring in the *Saturday Evening Post* under the title "Man in Chains" and presages the general content of the book. Two large classes of people form the author's main concern: those who, in effect, believe that 'by taking thought they can add a cubit to their stature' and those who blame their failure upon everything but themselves. With examples from case histories, he proceeds to demonstrate that there are no short cuts to the abundant life—that any solution may only be found through individual development of disciplined habit patterns. His plea is for the discovery and codification of the axioms of personality; "this book represents a step in that direction."

The author's prestige and forceful style should make this book effective in combatting the insidious paternalism with which some of our educational and governmental institutions are honeycombed. Employment counsellors who have not arrived at a stage of dangerous sophistification or become imbued with defeatism should find it valuable.

University of Maryland ALAN M. KERSHNER

A History of Philosophy. By B. A. G. FULLER. New York, Henry Holt & Co., 1938. Vol. I, pp. ix, 417; Vol. II, pp. 675.

Professor Fuller's *History of Philosophy* is one of the most readable ever written. Without pretending to add to the scholarship of the subject, it covers, in a vivid, modern, and, at times, intimate and racy style, twenty-five centuries of culture. Its graphic figures of speech and numerous analogies are drawn from all branches of knowledge and all walks of life.

This work does not contain formal sketches of the philosophers, which characterize many histories of philosophy, and footnotes and references are very rare. The few quotations that it does contain are not specific and refer to a volume rather than to a definite page. Yet the work is immense in size. At the end of the second volume there is appended for the convenience of the student a brief bibliography, a chronology, and a glossary of terms based on Lalande's *Vocabulaire de la Philosophie.*

While the author admits his leanings towards naturalism and materialism, he had no object in mind beyond writing an interesting history of philosophy. There is no attempt to treat thinkers out of their historical sequence in order to give a more logical development of the history of philosophy. Continuity is achieved by giving connecting links between thinkers and by introducing the thought of lesser philosophers, like Leonardo Da Vinci and Darwin, whose influence is usually relegated to footnotes or only casually mentioned and not featured as an essential part of the text.

The first five chapters, which are a condensation of the author's three volume *History of Greek Philosophy* are the most scholarly, and the section on medieval philosophy extending from the patristic philosophers to the fall of scholasticism is disproportionately brief. On the other hand, the accounts of Plotinus and Schopenhauer, which the author took over from his articles on the two men in the *Personalist,* are longer and fuller than is justified.

In addition to these lacks of balance, the work suffers from too many concessions to popular taste in style. For the psychologist who wants a readable history of philosophy rather than a handy reference book, this volume has, however, no equal.

Auburn, N.Y. SALVATORE RUSSO

L'angoisse et l'emotion. By RENÉ LACROZE. Paris, Boivin & Cie., 1938. Pp. x, 291.

This theoretical discussion of the evolutionary growth of the affective life presupposes that anguish, pain, or suffering furnishes the starting point for the whole range of feeling. Emotion is the generic name for the essential characteristic of this type of experience which develops into a host of variable phases depending upon the interplay of organic movements, the infiltration of ideas, and the expressions of the entire personality toward many environmental situations. Conflicts arouse many varieties of such emotional responses and thus occasion the criteria by means of which these experiences are practically or systematically classified on a qualitative basis, such as joys, disgusts, sadnesses and the like. In the ascending development of nervous patterns the discharge of ever more complicated, sometimes playful, activities are

made possible. Finally the ideational functions provide further patterns of emotional experience, making new delineations permissible.

The book is provocative of critical thought and shows a comprehensive grasp of the literature on the subject. Its thesis deserves careful and serious consideration. The text is well documented. It is followed by an extended bibliography and an index of authors cited. The first part of the treatise deals with the pattern or structure of emotion; the second with the lesser or subordinate types of affective experience including the abnormal manifestations. The doctrine advanced by the author goes one step beyond the reviewer's phylogenetic theory in that vague uncomfortableness alone accounts for the beginning of the affective life, instead of starting with both pleasantness and unpleasantness as the sources.

Chicago, Ill. CHRISTIAN A. RUCKMICK

Guiding Human Misfits. By ALEXANDRA ADLER. New York, Macmillan Co., 1938. Pp. viii, 88.

To some readers, the sub-title of this book, "A practical application of individual psychology," will seem more fitting than the main title. Eighty-eight pages, although direct, condensed, and well written, is all too few pages in which to deal with the enormous social and psychological problem of guiding human misfits. The author has confined her efforts mostly to presenting material which she obtained while observing and treating patients. No attempt is made to review the abundant material upon the subject.

The book is written from a modified psychoanalytic viewpoint; the author duly acknowledging her indebtedness to her father, Alfred Adler, for much of the theory and methodology. Having studied neurology, psychiatry, and psychotherapy she has obviously had a rich background of practical experience from which to draw illustrative material.

Emphasis is placed upon the significance of early experiences and training. A few case histories are given to illustrate how neurotic tendencies in childhood may ultimately lead to definite neurotic and psychotic conditions in later life. Possible preventives and treatments are suggested.

While many psychologists will question the validity of the general approach, and consequently the interpretations and conclusions, yet, it is believed that even these will admit considerable sound judgment is indicated in diagnosing and treating individual cases. The book is apparently intended for practical use by those working with misfits. It is readily comprehensible to the layman and will undoubtedly prove of interest to a wide variety of people.

Berea College LAWRENCE M. BAKER

Mental Conflicts and Personality. By MANDEL SHERMAN. New York, Longmans, Green & Co., 1938. Pp. vii, 319.

In this book, Dr. Sherman brings to a focus upon the understanding of mental conflicts the points of view of psychology, psychiatry and sociology. Psychiatrists and psychologists have, in the past, frequently exchanged ideas upon this subject, but the views of the medically trained person have far too often been widely separated from those of the sociologist. The result has been that social and cultural factors in adjustment have tended to be lost sight of in psychiatric interpretation. This book

goes far to remedy the defect. Due regard is given to the information which psychology has to offer, to the formulations of psychopathologists of various schools, and to the work of sociologists.

The study commences with a consideration of the genesis of mental conflicts. Both the innate drives and those of secondary origin receive recognition. The relation of attitudes to conflicts is then discussed. Next comes a more detailed study of the three main types of conflict. It is noteworthy that conflicts due to readjustment to new and varying cultural patterns are first considered, followed by conflicts of inferiority and insecurity, and eventually by sex conflicts. The last portion of the book relates the material concerning conflicts to, first, neuroses and, secondly, antisocial behavior in children. The balance between the three types of conflicts is maintained here, without the tendency to base every manifestation upon a single conceptual foundation, a tendency which is unfortunately far too common among those who deal with the complexities of human behavior.

Dr. Sherman has given us a lucid, readable book. While it contains no startlingly new ideas, it does present the material in a valuable new setting. His interest in children keeps him from going to the extreme of dealing only with adult manifestations and also from treatment the early period of life merely as the psychologist or psychiatrist wishes to remember it in an attempted projection of himself into it.

Milwaukee County Mental Hygiene Clinic GILBERT J. RICH

Introduction to Methods in Experimental Psychology. By MILES A. TINKER and KENNETH H. BAKER. New York, D. Appleton-Century Co., 1938. Pp. 222.

This is a spiral-ring-bound laboratory manual in psychology. Besides being bound in this inexpensive loose-leaf form the pages are perforated and punched for binding in the student's notebook. The material covered is modern, and is not limited to experimental psychology in the older sense but covers psychophysics, learning, including one experiment on animal learning, demonstration of the presence of various psychological traits such as "The Development of Concepts," "Estimation of Temporal Intervals," and "Measurement of General Ability." There are several experiments which are intended to bring out methods of statistical treatment of laboratory observations. The student is shown how to organize group data, to apply measures of variability, and there is some discussion of measurements of central tendencies, and other statistical material. The last eleven experiments deal with such topics and studies as "Musical Aptitude," "Reading Ability," "Art Appreciation," "Mechanical Ability," "Manual Dexterity," "Free Association," "Diagnostic Association," "Adjustment Inventory," "Self-Rating of Personal Traits," and "Individual Diagnosis by Means of the Profile." There are three appendices dealing with the terms and data needed for some of the experiments, and also some statistical tables. The last appendix consists of pictures for the experiment in "Judging Intelligence from Photographs." All the material needed for the study is present in this manual: the recording sheets, including cross-section paper, printed forms for tables, and other material for graphic and tabular recording are all present so that the student does not have to go outside of the manual for the materials in a course in experimental psychology.

Recorder's Court, Detroit, Michigan LOWELL S. SELLING

An Evaluation of Visual Factors in Reading. By H. A. ISMUS, J. W. M. ROTHNEY, and R. M. BEAR. Hanover, N.H., Dartmouth College Publications, 1938. Pp. 144.

This monograph is a description of the diagnostic and remedial reading program at Dartmouth College. In Chapter I are stated the specific problems to be investigated. Chapter II gives an exact description of the eye examinations, the procedure with the eye movement photography, the aptitude and reading tests, and the criterion of academic success that were used. The interrelationships between the various tests are given in Chapter III. An excellent critical appraisal of the Opthalm-O-Graph is given in Chapter IV. The reliability of the eye movement photography is studied, as are the relationships with the other variables considered. The plan and instructional procedure are given in Chapter V together with an evaluation of the results. In a final chapter a summary is given and conclusions are cautiously drawn.

The mass of data which was obtained is clearly presented in a liberal number of figures and tables. The special tests that were used are given in an appendix, together with the more extensive tables.

This study can well serve as a model for programs concerned with the experimental evaluation of diagnostic techniques and methods in remedial reading. The completeness with which the authors present their data will be of great aid in helping other workers plan similar projects. The whole study is remarkable for the extreme care taken in working out methods and procedures, the completeness with which the data were analyzed, and the unwillingness to generalize beyond the findings.

University of Maryland — EDWIN E. GHISELLI

—*And the Stutterer Talked.* By A. HERBERT KANTER and A. S. KOHN. New York, N.Y., Bruce Humphries, 1938. Pp. 236.

Cast in the form of fiction, this book was written to and for the layman. The hero goes through many struggles in his search for normal speech. The origin of several cases of stuttering are also recounted with their course of development through unwise parental tutoring, ridicule from companions and teachers, and treatments at the hands of numerous and persistent quacks.

Stuttering is defined as hesitative repetitive speech characterized by clonic or tonic spasms. It is emphasized that stuttering is a psychological phenomenon; a manifestation of obsessions and inhibitions. Stammering, with which the book is not concerned, is defined as speech impairment due entirely to anatomical or organic abnormalities of the speech mechanism. This distinction, which is not usually recognized by psychologists nor by the Warren *Dictionary,* is held to steadfastly and perhaps has its value.

While frequently referring to the importance of a defective inherited nervous constitution, which is claimed to be characteristic of the stutterer, much stress is given to the significance of learning and habits. "Speech is an acquired attribute, never inherited . . . the mental aspects of the condition are acquired step by step through the years" (p. 177). It is said that an obsessive state of mind has been reached which resulted from a series of accumulated bad speech experiences, and "Cure depends solely upon the extent to which the obsession can be eradicated. This means complete rehabilitation of the individual" (p. 176).

In general, the book fulfills quite well the purpose of focusing attention upon the

prevalence, the importance, and the lack of dependable treatment available to stutterers. Because there are no references, no indication of the amount or quality of the research upon which the conclusions are based, and a failure to isolate more adequately the various factors which presumably led to the successful cure, the book has limited scientific value for psychologists.

The authors were not attempting a scientific treatment, however, and judged as a popular book written upon an involved psychological subject, it is more acceptable than most of them.

Berea College — L. M. BAKER

The Study of Human Abilities. By J. K. SHRYOCK. New Haven, American Oriental Society, 1937. Pp. x, 168.

A translation and critical analysis of a book dealing with human nature and human abilities, written between 240 and 250 A.D. by a retired Chinese official named Liu Shao, should be of interest to modern students of human nature. The Chinese title is the *Jen Wu Chih,* and the author presents the materials with which it deals in the preface. He says, "The sages admired wisdom more than all else. Of all wisdom, there is nothing more valuable than to know men, and those who know men are certainly wise." The title of *Jen Wu Chih* as translated is *The Study of Human Abilities.* This manuscript is in reality an analysis of the text through a translation, interpretation and application of the contents in the light of both the period in which it was written as well as present day knowledge.

"The *Jen Wu Chih* is primarily a work on the applied psychology of character. It deals with the abilities of men. These abilities are mental traits which can be observed only through their manifestations in behavior and Liu Shao might be called a behaviorist" (p. 39). It is such a treatise that Shyrock makes of this study. The book begins with a study of the history and philosophy of the three kingdoms, followed by a short biography of Liu Shao. The history and characteristics of the materials are given some consideration, followed by a presentation of the philosophy, morals, science, psychology and conceptions of the state of Liu Shao as revealed in his writings. The analysis of *Jen Wu Chih* makes up the major portion of the book, and it is this part that the reader may find repetitious, and at points confusing, due to the continuous analysis and classification of all human abilities and characteristics. These materials reveal, however, the nature that Liu Shao used in analyzing man. The book would be more interesting to the reader if some of these materials had been presented more generally in a descriptive manner, rather than as definitions or in an analytic fashion. This study should certainly serve to give a more complete insight into the philosophy of this period in Chinese History.

North Carolina State College — K. C. GARRISON

Introduction to General Psychology. By JAMES B. STROUD. New York, Prentice-Hall, 1938. Pp. xv, 681.

In a day of popularizations, this general textbook is distinguished by the inclusion of basic material of a type that the fashionable text often omits. Thus, chapters on the neuro-muscular system, sensory processes, work, fatigue, sleep, and pleasantness and unpleasantness, present experimental results which appear, refreshingly enough, to have been selected for educational rather than entertainment value.

It is unfortunate, therefore, that the potential usefulness of a presentation of a

wide range of information has been impaired by a lack of clarity and organization. One may seriously question whether a student could obtain a balanced and informed concept of the general field of psychology from data which are often scattered through the pages with little attempt at interpretation and analysis. While the selection of many of the topics suggests a functionalistic bias on the part of the author, the tendency never serves to integrate the great mass of material. When interpretations are offered they too often consist in naming a theory or viewpoint and presenting a somewhat hazy portrait of its proponent. Other interpretations are sometimes proffered in complete isolation from the remainder of the material, a procedure which leads to the inconsistent use of undefined terms.

In short, this attempt has more of the characteristics of a handbook than of a general text. Its usefulness as a handbook is destroyed, however, by the frequent employment of secondary sources, and by many misinterpretations and murkily written sections.

The quality of the printing, binding, and indexing of the volume is consistent with the carelessness of its written construction.

Tulane University S. RAINS WALLACE

Kleine Logik der Geisteswissenschaften. By FRANZ SCHMIDT, Munich, Ernst Reinhardt, 1938. Pp. 128.

Those who have followed the philosophical development of German science before the present "extirpation of the German mind in favor of the German race" (a development predicted by Nietzsche in 1871) must regard this little book as a tragic rudiment of a great work. It is concerned with the logic of the mental sciences—the development of which started with Hoeffding and Dilthey. Schmidt's book is a survey of the persisting problems; it offers little that is new. Maintaining that psychology has not fulfilled the promise to give a basic interpretation of the mental sciences, Schmidt calls upon logic and tries to find the basic principles in the notion of co-intention, a concept which goes back to Husserl's intentional philosophy. The discussion is divided into three parts: "The Mental Experience," "The Order of the Mental World," and "The Border of the Mental World." Of most interest to the psychologist is Schmidt's logical interpretation of the psychological types and of the rôle of typology (p. 58). The most original part of the book is his interpretation of freedom (p. 96).

Washington, D.C. ERNST HARMS

Psychology and the Nurse. By FRANK J. O'HARA. Philadelphia, W. B. Saunders Co., 1939. Pp. 252.

This book contains material touching upon the sense organs and sensation, dreams, memory, emotion, conflicts, the art of thinking, attention, tests and measurements, personality, and mental disorders. The writing is simple and the analysis superficial. Applications of each topic to nursing are discussed and a section at the end of each chapter is devoted to questions for review.

The author tends to use the definite article too freely, to devise questionable terms, to use others in a slipshod manner, to employ obsolete categories, and in general to oversimplify the discussion.

University of Maryland LESTER P. GUEST